RESOURCES *for the* FUTURE

Welfare Losses from Wildfire Smoke: Evidence from Daily Outdoor Recreation Data

Jacob Gellman, Margaret Walls, and Matthew Wibbenmeyer

Working Paper 23-31
August 2023

About the Authors

Jacob Gellman is a Postdoctoral Fellow in the Department of Economics at the University of Alaska Anchorage. His research focuses on the economics of wildfire, land use, and housing. He holds a PhD in Environmental Science & Management from the University of California, Santa Barbara. Throughout his career he has worked on interdisciplinary wildfire issues with economists, ecologists, meteorologists, and other natural scientists. Prior to his graduate studies, he worked as an energy economics consultant, where he advised utilities and tribes on energy decisions and produced expert witness testimony for legal cases.

Margaret Walls is a senior fellow at Resources for the Future (RFF). Her current research focuses on issues related to resilience and adaptation to extreme events, ecosystem services, and conservation, parks and public lands. Walls's work on resilience assesses the factors that affect household location decisions in coastal areas, how individuals perceive flood risks, and how risk perceptions affect adaptation decisions. She has estimated the value of natural lands—such as wetlands—in providing protection from hurricanes and flooding and is assessing the extent to which hurricanes affect US migration patterns.

Matthew Wibbenmeyer is a fellow at RFF. His research studies climate impacts and mitigation within the US land sector, with a special emphasis on wildfire impacts and management. US wildfire activity has accelerated in recent years, leading to increases in property damages, carbon emissions, and health impacts due to smoke. Wibbenmeyer's research studies the impacts of these changes for communities, how these impacts are distributed, and how management choices affect the distribution of impacts. Alongside his work on wildfire, Wibbenmeyer is investigating the role of the US land sector in mitigating climate change, and how policy toward land sector choices may influence the United States' ability to meet climate goals.

Acknowledgments

We thank Andrew Plantinga, Olivier Deschênes, Kelsey Jack, Max Moritz, Dave Lewis, Steve Dundas, Jeff Englin, and Molly Robertson for comments and suggestions on this paper. We are also grateful to seminar participants in the UCSB Environmental Economics group, Western Economics Association 2022 Conference, AERE 2021 Summer Conference, 2023 Occasional Workshop, AERE OSWEET, and University of Nevada-Reno. In addition, we thank Jude Bayham and coauthors, who generously shared the gridded air quality data. This research was funded by a United States Department of Agriculture National Institute of Food and Agriculture (NIFA) Agriculture and Food Research Initiative (AFRI) grant, award number 2020-67023-33258.

About RFF

Resources for the Future (RFF) is an independent, nonprofit research institution in Washington, DC. Its mission is to improve environmental, energy, and natural resource decisions through impartial economic research and policy engagement. RFF is committed to being the most widely trusted source of research insights and policy solutions leading to a healthy environment and a thriving economy.

Working papers are research materials circulated by their authors for purposes of information and discussion. They have not necessarily undergone formal peer review. The views expressed here are those of the individual authors and may differ from those of other RFF experts, its officers, or its directors.

Sharing Our Work

Welfare Losses from Wildfire Smoke: Evidence from Daily Outdoor Recreation Data*

Jacob Gellman[†], Margaret Walls[‡], Matthew Wibbenmeyer[§]

July 7, 2023

Abstract

Wildfire smoke pollution is growing in the western United States. Estimates of its health impacts are numerous, but few revealed preference estimates of its damages exist. We study a setting where individuals are directly exposed to smoke, and avoidance behavior is measured with high frequency: outdoor recreation. We combine millions of administrative campground reservation records with satellite data on wildfire, smoke, and air pollution. These data are rich among most studies of recreation, with nearly 1,000 campgrounds and detailed individual-level observations. The data allow us to model sequential recreation decisions under evolving information using a novel control function approach. We estimate that wildfire smoke reduces welfare by $107 per person per trip. Damages are larger when campgrounds are affected by consecutive days of smoke and attenuated when smoke events are sufficiently far from active fires. In total, 21.5 million outdoor recreation visits in the western United States are affected by wildfire smoke every year, with annual welfare losses of approximately $2.3 billion. These findings contribute to a growing body of evidence on the costs of wildfire smoke.

Keywords: Wildfire smoke, wildfire, air pollution, recreation
JEL codes: Q26, Q51, Q53, Q54

*We thank Andrew Plantinga, Olivier Deschênes, Kelsey Jack, Max Moritz, Dave Lewis, Steve Dundas, Jeff Englin, and Molly Robertson for comments and suggestions on this paper. We are also grateful to seminar participants in the UCSB Environmental Economics group, Western Economics Association 2022 Conference, AERE 2021 Summer Conference, 2023 Occasional Workshop, AERE OSWEET, and University of Nevada-Reno. In addition, we thank Jude Bayham and coauthors, who generously shared the gridded air quality data. This research was funded by a United States Department of Agriculture National Institute of Food and Agriculture (NIFA) Agriculture and Food Research Initiative (AFRI) grant, award number 2020-67023-33258.

†University of Alaska Anchorage. Email: jgellman@alaska.edu.
‡Resources for the Future, Washington, DC. Email: walls@rff.org.
§Resources for the Future, Washington, DC. Email: wibbenmeyer@rff.org.

1 Introduction

Large wildfires have increased in frequency and severity in the western United States, and these trends are expected to continue as the climate warms (Abatzoglou and Williams 2016; Westerling 2016, 2018; Williams et al. 2019). This wildfire activity has generated considerable increases in smoke, which can cover large geographic areas and affect air quality hundreds of miles away. Wildfire smoke now accounts for up to half of particulate matter pollution in some areas of the western United States (Burke et al. 2021), and some literature has found that its particulate emissions harm health more severely than particulate emissions from other sources (Aguilera et al. 2021; Kochi et al. 2010). Empirical evidence on the negative effects of wildfire smoke is growing, with studies finding increased morbidity and mortality, negative effects on birth outcomes, reductions in academic performance, and impacts in labor markets (Cullen 2020; Heft-Neal et al. 2023; Kochi et al. 2010; McCoy and Zhao 2021; Miller et al. 2021; Reid et al. 2016; Wen and Burke 2022; Borgschulte et al. 2022).

Few estimates of the welfare costs of smoke exist, however, and those that do are based on stated preference or survey-based methods (Richardson et al. 2012, 2013; Jones 2017) or were calculated by applying a value of statistical life (VSL) to changes in mortality (Miller et al. 2021). Welfare estimates derived from revealed preference methods are limited, perhaps because smoke is challenging to study in a revealed preference setting. In hedonic property value studies, it may be difficult to distinguish effects of long-run increases in wildfire smoke from other unobserved region-level changes to housing markets. In travel cost recreation studies, data at a high temporal resolution are needed to compare changes in behavior in response to transient smoke conditions.

We use high-frequency data on outdoor recreation in a travel cost framework to provide the first revealed preference estimates of welfare damages from wildfire smoke. Outdoor recreation is advantageous for studying responses to smoke for several reasons. First, exposure to smoke is high for recreationists. Wildfire season and peak outdoor recreation season tend to coincide, with more than 1 million national park visitor-days per year taking place during hazardous smoke conditions (Gellman et al., 2022). Second, visitors to recreation sites spend large amounts of time outdoors and tend to engage in vigorous activities, such as hiking, which may exacerbate the effects of exposure (Korrick et al. 1998; Richardson et al. 2012). Smoke can also reduce the visibility and amenity value from recreation trips. Third, structural modeling of outdoor recreation decisions using the travel cost framework can naturally yield welfare cost estimates (Parsons 2017; Lupi et al. 2020).

Our approach uses millions of administrative campground reservation records from nearly 1,000 federally managed campgrounds, which we combine with daily data on wildfire, smoke, and air pollution. The data include information on campsite reservations, which are typically made well in advance, and cancellations of planned trips. The cancellations data allow us to account for the short-term and transient nature of wildfire smoke by focusing on decisions made both before and after visitors have knowledge of air quality conditions at the site.

We model visitors' reservation and cancellation decisions as a two-stage discrete choice. In the first stage, a visitor chooses to reserve based on expected site conditions; in the second stage, they decide whether to cancel based on realized site conditions. A key feature of this setting is that reservations and cancellations are made in sequence; a visitor can only cancel if they have a reservation, meaning they have already demonstrated a preference for the site. This induces sample selection in the population holding reservations; if unaccounted for, this would bias welfare estimates of wildfire smoke. To correct for this bias, we develop a novel control function approach to link preferences across choices (Wooldridge 2015). The control function uses estimated preferences from the first-stage reservation decision to remove selection bias in the second-stage cancellation decision. We demonstrate the effectiveness of this approach through numerical simulations.

We find that wildfire smoke imposes welfare costs of \$107 per person per recreation trip. Without accounting for sample selection using the control function, the analysis would have implied damages of \$154 per person per trip, overstating welfare impacts by 44 percent. Damage estimates vary by the duration of smoke events, as measured by the number of smoke-affected days in the week leading up to arrival. When a campground is affected by smoke on only one day in the week of arrival, estimated damages are as low as \$32 per person per trip; but when it is affected on all seven days, losses are as high as \$432 per person per trip. The heterogeneity in our estimates is consistent with greater perceived likelihood of experiencing smoke during weeks with more frequent smoke, more severe smoke impacts, or potentially both. In the appendix, we report additional heterogeneity and robustness checks, including a placebo for smoke, heterogeneous responses by distance to an active wildfire, and heterogeneous responses by type of recreation site.

The overall magnitude of welfare losses from wildfire smoke is large. We find that in the western United States, an average of 21.5 million outdoor recreation visits per year (approximately 4.2 percent of all annual visits), which includes both camping and noncamping trips, are affected by wildfire smoke on lands managed by the National Park Service, US Forest Service, Bureau of Land Management, US Army Corps of Engineers, and states. Applying the empirical welfare estimate of \$107 per person per trip, this figure implies welfare losses of roughly \$2.3 billion per year.

Our study adds to the growing literature on wildfire and wildfire smoke. Several studies have examined short-run consequences of smoke, finding that it increases emergency room visits (Heft-Neal et al. 2023) and mortality (Miller et al. 2021), decreases labor earnings (Borgschulte et al. 2022), and worsens expressed sentiment on social media (Burke et al. 2022; Loureiro et al. 2022). Some researchers have used survey evidence to estimate welfare impacts of smoke. Based on US life satisfaction survey data, for example, Jones (2017) estimates a willingness to pay (WTP) of \$373 to avoid one wildfire smoke day per six-month period. Richardson et al. (2012) employ a defensive expenditure approach using survey data on air purifier purchases to estimate a WTP to reduce one smoke-induced health symptom day of \$84 (in 2009 dollars). Using a stated preference contingent valuation survey, Richardson et al. (2013) estimate a WTP to avoid one smoke-induced symptom day of \$95, which they compare to an \$87 WTP estimate based on a cost-of-illness approach (both figures in 2009 dollars).

We measure the welfare damages of smoke exposure using cancellations of planned recreation trips; our estimate of \$107 per trip roughly translates to \$38 per day. This value is in line with estimates from Richardson et al. (2012, 2013) but considerably lower than the estimate from Jones (2017). Compared to our aggregate welfare estimate of \$2.3 billion per year, Miller et al. (2021) value damages from mortality among elderly Medicare recipients in the United States at \$6–170 billion, depending on VSL assumptions. Borgschulte et al.

(2022) find annual lost labor earnings of $125 billion per year due to wildfire smoke.

Measuring the cost of wildfire smoke is crucial to informing public policy. The federal government spent an average of $2.8 billion per year on fire suppression between 2017 and 2021, and the State of California spent an average of $900 million per year from 2018 to 2022.[1,2] California has proposed spending $1.2 billion over fiscal years 2022–23 and 2023–24 on wildfire mitigation measures, including vegetation management, prescribed burns, and home hardening.[3] These activities are consistent with the state's recently declared goal to treat 1 million acres of hazardous fuels per year.[4] Together, the 2021 Infrastructure Investment and Jobs Act (P.L. 117-58) and the Inflation Reduction Act of 2022 (P.L. 117-169) provided $24 billion for federal wildfire programs (Goldman et al. 2022). Understanding the welfare costs of wildfires and smoke is critical to assessing the benefit of these public expenditures.

We also contribute to the recreation demand literature by using novel methods and data sources. We value a transient environmental disamenity in a setting where users make decisions under evolving sets of information. The two-stage choice structure that links preferences across decisions is informed by literature on sample selection correction in nonlinear models (Greene 2012; Terza 2009) and recreation contexts (Cameron and DeShazo 2013; Cameron and Kolstoe 2022; Kolstoe and Cameron 2017; Lewis et al. 2019). Our framework could be used to model sample selection or sequential choices in other nonlinear or discrete-choice settings. It could also be applied to recreation studies valuing other short-lived environmental amenities, such as temperature and precipitation extremes or acute pollution events. In addition to the modeling, our use of administrative data contributes to a recent literature using new, large, or innovative data to study recreation across broad regions (Cameron and Kolstoe 2022; Dundas and von Haefen 2020; English et al. 2018; Parthum and Christensen

[1] National Interagency Fire Center. Suppression Costs. `https://www.nifc.gov/fire-information/statistics/suppression-costs`.

[2] California Department of Forestry and Fire Protection. Suppression Costs. `https://www.fire.ca.gov/stats-events`.

[3] California Legislative Analyst's Office. The 2022–23 Budget Wildfire and Forest Resilience Package. `https://lao.ca.gov/Publications/Report/4495`.

[4] Agreement for shared stewardship of California's forest and rangelands between the State of California and the USDA Forest Service, Pacific Southwest Region. `https://www.gov.ca.gov/wp-content/uploads/2020/08/8.12.20-CA-Shared-Stewardship-MOU.pdf`.

2022).

The remainder of this paper is organized as follows. In Section 2, we describe the data sources for the study, including recreation, smoke, fire, and pollution data. We also discuss several descriptive features of the data. Section 3 describes the modeling approach, including a conceptual framework and a description of the estimating dataset. In Section 4, we turn to estimation, describing various sets of results. Section 5 appraises the total annual welfare damages of wildfire smoke in the west. Section 6 concludes.

2 Data

We combine data on recreation behavior, travel costs, wildfire smoke, air pollution, wildfire activity, and weather into three main datasets. The first is a daily panel of federally managed campgrounds in the western United States, 2010–2017, that includes climate normals and local daily measurements of smoke, wildfire activity, pollution, and weather. The second dataset is a record of individual-level reservations for campgrounds, which we link to the daily campground panel to show site conditions for users' reservation dates. The last dataset aggregates the individual users to measure daily reservation activity by distance from each campground.

2.1 Recreation

We obtained data on campground use from Recreation.gov,[5] the primary web portal for reserving sites at federally managed campgrounds, including those managed by the National Park Service, Bureau of Land Management, US Forest Service, US Army Corps of Engineers, and Bureau of Reclamation. Figure A1 in Appendix A displays the web interface as a user would experience it. The website gives users information about campground amenities, prices, availability, and nearby points of interest.

The raw data include more than 90 million transactions from more than 7 million unique users. We limit attention to campgrounds in the 11 western states, during the months of May–September, and for the years 2010–2017, which leaves more than 16 million transactions

[5]Recreation.gov. `https://www.recreation.gov`.

from 2 million unique users at 999 campgrounds. Our analysis is primarily concerned with overnight camping and excludes, for instance, large-group or equestrian facilities.

The data give detailed information on reservations, walk-ins, cancellations, no-shows, transaction dates, payments, refunds, zip code of origin, group size, user identifiers, and other information. For every transaction in an order, such as a payment or cancellation, the exact time is known. For the 999 campgrounds, 84 percent of transactions were made online, 9 percent over the phone, and 7 percent on-site (such as walk-ins or early checkouts).

2.2 Travel costs

We calculate travel costs based on the distance and travel time between a origin zip codes and destination campgrounds with GraphHopper, an open source routing engine that uses Djikstra's algorithm and OpenStreetMap data.[6,7] We calculate nearly 5.4 million routes representing 5,379 origin points and 999 destinations. Our estimates reflect the fastest routes by car between each origin and destination. Optimal routes generally match routes identified by Google Maps during periods of low traffic. To identify the coordinates of each user's zip code, we match zip codes to Census Zip Code Tabulation Areas (ZCTAs) and find the centroid of each ZCTA.[8,9] Figure A2 in Appendix A displays an example route.

Following English et al. (2018), we calculate the per-person travel costs between ZCTA z and campground j as

$$c_{zjt} = \frac{p^D_{zt} D_{zj}}{n} + p^T_{zt} T_{zj}, \tag{1}$$

for travel distance D_{zj} and travel time T_{zj}. The per-kilometer cost of traveling between ZCTA z and campground j is given by p^D_{zt} and includes gasoline costs, per-kilometer vehicle maintenance, and per-kilometer average vehicle depreciation. For gasoline costs, we use state- and year-specific averages of per-kilometer gasoline costs during summer months,

[6] GraphHopper. https://www.graphhopper.com.

[7] GraphHopper GitHub. https://github.com/crazycapivara/graphhopper-r.

[8] Health Resources and Services Administration, John Snow, Inc., and American Academy of Family Physicians. Uniform Data System. https://udsmapper.org/zip-code-to-zcta-crosswalk.

[9] Because ZCTA centroids may not be located along roads, we snapped them to the nearest road using Census TIGER/Line shapefiles and used the nearest points along roads as origin points.

based on per-gallon gasoline costs from the Energy Information Administration and nationwide average fleet fuel economy.[10,11] We use per-kilometer average depreciation and vehicle maintenance costs from AAA data, as in English et al. (2018).[12] Last, we measure hourly costs of travel time p^T_{zt} as one-third of the average household income in ZCTA z divided by 2,080 hours worked per year (English et al. 2018). All numbers are inflation-adjusted to 2020 US dollars.

2.3 Smoke and air pollution

Our measure of smoke impacts combines satellite-derived wildfire smoke plume data and gridded ground-level $PM_{2.5}$ monitoring data. For each day, we record whether a campground was covered by smoke based on daily observations of smoke plumes from the National Oceanic and Atmospheric Administration (NOAA) Hazard Mapping System (Schroeder et al. 2008).[13] Each day, NOAA analysts manually trace the perimeters of smoke plumes using satellite photography, producing daily shapefiles. These data have been used in studies examining the effect of smoke on air pollution, health, labor markets, self-protective behavior, and crime (Borgschulte et al. 2022; Burke et al. 2021, 2022; Burkhardt et al. 2019; Cullen 2020; Gellman et al. 2022; Heft-Neal et al. 2022; Miller et al. 2021; Preisler et al. 2015).

One challenge presented by this dataset is that satellite photography does not reveal where in the air column a smoke plume is: it could be at the ground level or high in the atmosphere. If the latter, it may not reflect on-the-ground conditions. To address this challenge, we code an area as smoke affected only if it is covered by a smoke plume and its ground-level $PM_{2.5}$ is above at least 1.64 standard deviations of the location-specific seasonal mean for nonsmoke days, which represents the 95th percentile of a normal distribution (Burkhardt et al. 2019; Gellman et al. 2022).[14] Figure A3 in Appendix A displays an example of that restriction using kriged $PM_{2.5}$ data from Burkhardt et al. (2019). The map

[10] Energy Information Administration. Weekly Retail Gasoline and Diesel Prices. https://www.eia.gov/dnav/pet/pet_pri_gnd_a_epmr_pte_dpgal_m.htm.

[11] Bureau of Transportation Statistics. Average Fuel Efficiency of US Light Duty Vehicles. https://www.bts.gov/content/average-fuel-efficiency-us-light-duty-vehicles.

[12] For example: AAA. Your Driving Costs 2016. https://publicaffairsresources.aaa.biz/wp-content/uploads/2016/03/2016-YDC-Brochure.pdf.

[13] NOAA. Hazard Mapping System. https://www.ospo.noaa.gov/Products/land/hms.html.

[14] A "season" is defined as fall, winter, spring, or summer.

shows that although many areas were covered by smoke, only some had air quality poor enough to be coded as smoke affected.

2.4 Wildfire activity

At each campground, we measure the daily distance to an actively burning fire. We combine NASA MODIS fire detection points with the United States Geological Survey Monitoring Trends in Burn Severity (MTBS) fire perimeter dataset.[15,16] The MODIS detection points record 1 km centroids of fire activity at a daily resolution, including agricultural and prescribed fires (Giglio et al. 2016). The MTBS data map the final perimeters for US wildfires. Combining these data has two advantages. First, the use of known wildfire perimeters filters out any MODIS points not associated with a large wildfire. Second, the MODIS detection points limit attention only to the portion of a wildfire that was burning on a given day. We use a 1 km buffer around the fire's final perimeter and its start and containment dates to filter MODIS points. Figure A5 in Appendix A demonstrates an example of this process for the western United States.

2.5 Temperature and precipitation

To control for weather conditions, we gather daily precipitation (mm) and maximum and minimum temperature (°C) for every campground. These data are published at a 4 km resolution by the PRISM Climate Group at Oregon State University.[17] In addition, at each campground, we record 30-year climate normals that reflect average conditions for 1980–2010.

2.6 Descriptive features of the data

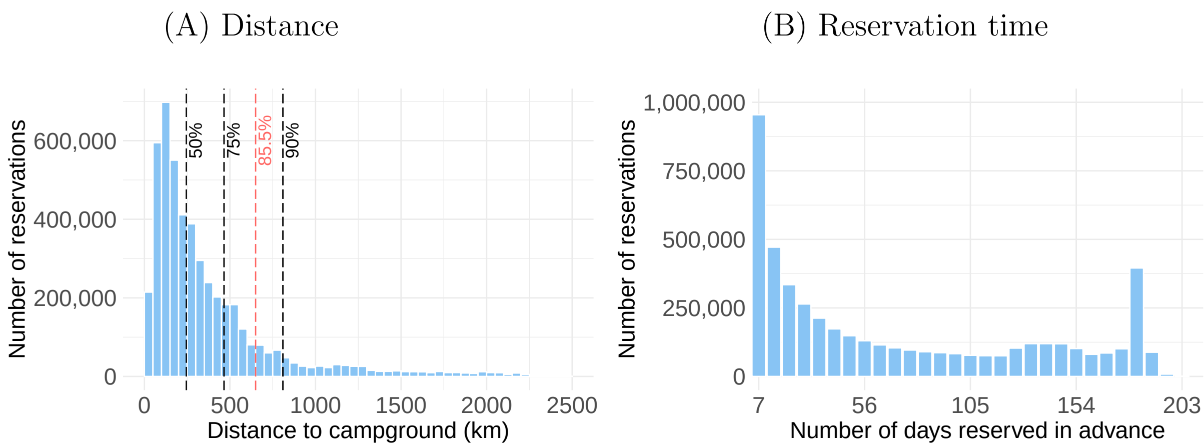
Of the 999 campgrounds in the analysis, 908 are managed by the US Forest Service, 50 by the National Park Service, 31 by the US Army Corps of Engineers, 5 by the Bureau of Land Management, and 5 by the Bureau of Reclamation; Figure A6 in Appendix A plots a map of

[15]NASA. Earthdata. https://earthdata.nasa.gov.

[16]USGS. Monitoring Trends in Burn Severity. https://www.mtbs.gov.

[17]Northwest Alliance for Computational Science and Engineering, Oregon State University. PRISM Climate Data. https://www.prism.oregonstate.edu.

Figure 1: Driving Distance and Time between Reservation and Scheduled Arrival



Note: Panel A shows the one-way driving distance of reservations from the destination campground, where the red line indicates a 650 km cutoff. Panel B displays the distribution of time reserved in advance of arrival date).

these. Although the Forest Service manages most of the campgrounds, the most-visited ones tend to be in national parks. Table A1 in Appendix A reports the most-visited campgrounds.

For the main analysis, we restrict the set of potential reservers to residents living within one day's driving distance of a given campground. We set this restriction at 650 km (400 miles). English et al. (2018) report survey results showing that, beyond 500 miles of driving distance, a substantial portion of recreation visitors are likely to have flown to their destination, which adds additional complications to the calculation of travel cost. Panel A of Figure 1 shows that our 650 km restriction results in including more than 85 percent of reservations in the dataset. Half of our observed trips come from within 250 km (155 miles), and three-quarters come from within 450 km (280 miles). Appendix H reports the main results using alternative driving distance thresholds.

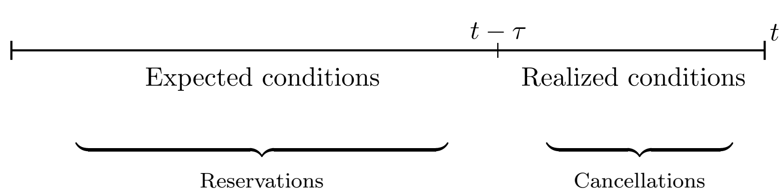
The timing of a reservation is also key for our setting. Wildfire smoke is a random event, meaning that visitors who reserve far in advance do not know whether their chosen campground will be smoke affected during their visit. Panel B of Figure 1 shows that most visitors reserve far in advance, consistent with results in Walls et al. (2018). Although a plurality of visitors reserve within a week of arrival, a majority reserve early. In addition, a significant mass appears around six months in advance, which is the earliest that some popular destinations allow reservations. In the following section, we describe our modeling approach to study the cancellation decisions of visitors who reserved ahead of time.

3 Modeling approach

In this section, we model an individual's decision to visit a campground under smoke and nonsmoke conditions. A key feature of this setting is that smoke is ephemeral; it is not a permanent feature of site quality but may impact air quality suddenly and without warning for several days or weeks. Given this, one estimation strategy might be to contrast reservation rates during smoky and nonsmoky periods.[18] However, because demand for campgrounds is high, campers must often make reservations far in advance, before smoke conditions are known. As campgrounds are often completely full shortly before a given date, rates of new reservations may not be responsive to smoke.

We therefore consider the cancellation decisions of visitors who reserved ahead of time. The population that is eligible to cancel consists of those that already hold a reservation; therefore, the sample is selected due to their demonstrated preference for the site. We account for selection bias induced by the reservation decision using a control function approach that explicitly models and estimates the "first-stage" reservation decision, in which visitors choose whether to reserve based on expected site conditions. In the second stage, close to the arrival date, they decide whether to cancel based on realized site conditions. Figure 3 illustrates the timing of these decisions, where t gives the arrival date and τ denotes a bandwidth sufficiently close to the arrival date. Our control function uses preference parameters estimated in the first stage to remove selection bias in the second stage.

Figure 3: Timing of Decisions



[18]For a discussion of late reservers, see Appendix B.

3.1 Reservations

Define utility from visiting site j at time t as

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} + \upsilon_{ijt}, \tag{2}$$

where V_{ijt} gives the observable portion of utility to household i from visiting campground j on date t, ε_{ijt} represents fixed unobserved preferences of household i for visiting site j at time t, and υ_{ijt} represents unobserved shocks to preferences that occur between time $t - \tau$ and time t. Let V_{ijt} be specified according to

$$V_{ijt} = \begin{cases} \delta c_{ijt} + \phi s_{jt} + X'_{jt}\gamma + \psi_j + \lambda_t, & j \in \{1, 2, ..., J\}; \\ 0, & j = 0, \end{cases} \tag{3}$$

where c_{ijt} gives the travel cost for person i to site j at time t, and s_{jt} represents smoke conditions at campground j on date t. The vector X_{jt} includes time-varying, campground-level conditions, such as precipitation, temperature, and proximity to an active wildfire. Additional variables include ψ_j and λ_t, which account for site- and time-specific fixed effects, respectively. Campground fixed effects capture time-invariant traits, such as quality or desirability. Time fixed effects, such as for year, week-of-year, and day-of-week, account for seasonality and trends in preferences over time. The parameter of interest is the WTP to avoid smoke, which is found by taking the ratio of marginal disutility of smoke ϕ to the marginal disutility of expenditure δ, WTP $= \phi/\delta$.

When making a reservation at an early date, person i knows their fixed unobserved preference for the site ε_{ijt} but not what site conditions or the shock υ_{ijt} will be. Therefore, they choose based on ε_{ijt} and an expectation of V_{ijt}; in expectation, υ_{ijt} is 0. Denoting the expectation of variable Y at the time of reservation as $\bar{Y}$, write the expectation of V_{it} as

$$\bar{V}_{ijt} = \begin{cases} \delta c_{ijt} + \phi \bar{s}_{jt} + \bar{X}'_{jt}\gamma + \psi_j + \lambda_t, & j \in \{1, 2, ..., J\}; \\ 0, & j = 0. \end{cases} \tag{4}$$

At the time of reservation, an individual's expected utility from visiting site j at time t is

$\bar{V}_{ijt} + \varepsilon_{ijt}$, where ε_{ijt} is known to the individual but unobserved by the econometrician. From the perspective of the individual, the expected value of υ_{ijt} is 0. Let $R_{ijt} = 1$ if an individual chooses to make a reservation to visit site j at time t. Individuals choose $R_{ijt} = 1$ if the expected utility of doing so exceeds the utility of their outside option, the observable portion of which we normalize to 0. Under the assumption that ε_{ijt} is distributed iid type I extreme value, the probability an individual makes a reservation is given by

$$\mathbb{P}(R_{ijt} = 1) = \mathbb{P}(\varepsilon_{i0t} - \varepsilon_{ijt} \leq \bar{V}_{ijt}) = \frac{\exp(\bar{V}_{ijt})}{1 + \exp(\bar{V}_{ijt})}. \tag{5}$$

For each campground j and day t, we sum the number of reservers and nonreservers in concentric zones z around a campground. The former (denoted N^1_{zjt}) is counted based on the reservations in the Recreation.gov dataset; for example, a reservation for four people is counted as four reservers. The nonreservers (N^0_{zjt}) are determined based on zip code–level populations within each concentric ring, less the number of people from each zip code that held a reservation to a different campground on that day. The unit of observation for the zonal estimation is a campground by day by 50 km distance bin, where each row of data reports the number of people choosing outcome variable $R_{ijt} \in \{0, 1\}$. Let ω denote the set of parameters $\{\delta, \phi, \gamma, \psi_j, \lambda_t\}$. Summing the individual log-likelihood function over all reservers and nonreservers in each zone and for each campground and date, the log-likelihood function is given by

$$\ell^R(\omega) = \sum_{z=1}^{Z}\sum_{j=0}^{J}\sum_{t=1}^{T} N^1_{zjt} \log\big(\mathbb{P}(R_{i(z),j,t} = 1|\omega)\big) + N^0_{zjt} \log\big(1 - \mathbb{P}(R_{i(z),j,t} = 1|\omega)\big), \tag{6}$$

for Z zones, T choice occasions, and J sites. Maximizing Equation 6 yields utility parameters for a representative individual i from zone z.

We use a zonal approach for the reservation decision for several reasons. The primary purpose of this first-stage reservation estimation is to construct a control function that accounts for preferences in the cancellation estimation. A person can only cancel a trip if they held a reservation. Therefore, preferences from the reservation decision likely play a role in the cancellation. The zonal reservation model accounts for these preferences and provides computational advantages over a multinomial logit approach. In this setting, we have more than 1,200 arrival dates to define choice occasions, nearly 5 million reservations, nearly 5,400 zip codes to account for the nonreserving individuals, and 999 campgrounds to form the choice set. It would be infeasible to use all the data in a multinomial logit model of individual site choice. We could reduce the size of the dataset by, for example, restricting the study to a single region or year. However, smoke is temporally and spatially correlated within regions, meaning that we require multiple regions and years to provide necessary variation and that site substitution is less likely to play a role in identifying the smoke parameter (see Appendix C). Because we require regional and temporal variation, fixed effects are crucial to remove location- and time-specific unobservables across many heterogeneous sites. The zonal model accommodates a high number of fixed effects and is computationally less expensive than the contraction mapping method used in many multinomial logit studies (Berry 1994). This computational speed makes a difference when bootstrapping standard errors in the two-stage model.

3.2 Cancellations

For the second-stage cancellation decision, we model a binary choice at the level of the individual trip.[19] At time $t-\tau$, site conditions are realized and approximately known to the individual, resulting in random shock $e_{ijt} \equiv V_{ijt} - \bar{V}_{ijt} + \upsilon_{ijt}$ to the utility of site visitation. The shock e_{ijt} includes realization of the actual (V_{ijt}) as opposed to mean ($\bar{V}_{ijt}$) smoke and other weather conditions and a random idiosyncratic and individual-specific taste shock (e.g., illness) represented by υ_{ijt}. Like ε_{ijt}, we assume υ_{ijt} is distributed type I extreme value. However, we allow its standard deviation to differ from the standard deviation of ε_{ijt} by a scale factor of $\frac{1}{\rho}$, letting $\upsilon_{ijt} \equiv \frac{1}{\rho}\eta_{ijt}$ where η_{ijt} is distributed standard type I extreme value.

[19] An alternative would have been to model a choice to cancel and rebook at another campground. In Appendix C, we show that very few users do so for the same choice occasion. Close to the arrival date, many campgrounds are fully booked, which can prevent site substitution. In addition, because smoke conditions are spatially and temporally correlated among potential alternatives, substitution is unlikely to be an important factor in the identification of the smoke parameter, given low variation in smoke conditions.

Conditional on having a reservation, individual i follows through on their reservation if

$$\bar{V}_{ijt} + \varepsilon_{ijt} + e_{ijt} \geq \varepsilon_{i0t} + e_{i0t}. \tag{7}$$

Using the definition of e_{ijt}, and letting C_{ijt} be a binary variable equal to 1 if individual i cancels their reservation within τ days of arrival, the probability that i follows through is

$$\begin{aligned}\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1) &= \mathbb{P}(V_{ijt} + \varepsilon_{ijt} + \upsilon_{ijt} \geq \varepsilon_{i0t} + \upsilon_{i0t}) \\ &= \mathbb{P}(\eta_{i0t} - \eta_{ijt} \leq \dot{V}_{ijt} - \rho(\varepsilon_{i0t} - \varepsilon_{ijt})),\end{aligned} \tag{8}$$

where the second line follows from the definition of υ_{ijt} and where $\dot{V}_{ijt} = \rho V_{ijt}$, a rescaling.

Equation 8 presents challenges for the econometrician. The variables ε_{i0t} and ε_{ijt} are unobserved. However, omitting these variables will bias parameter estimates because they are correlated with V_{ijt}, as only households with high taste for the site (ε_{ijt}) will have made a reservation. Specifically, ε_{ijt} is correlated with travel cost in the selected sample; for individuals with a high travel cost c_{ijt}, selection into the group of reservers implies a higher ε_{ijt}. Without sample selection correction, this relationship downward biases estimates of the travel cost parameter δ in the cancellation decision and thus inflates estimates of WTP $= \phi/\delta$. Section 3.4 explores this relationship using a numerical example. We show that the bias arises only when unobserved preferences from the first stage affect the second-stage decision ($\varepsilon_{i0t} - \varepsilon_{ijt} \neq 0$ in Equation 8) and when we can only observe the cancellation decision for the selected sample of reservers ($R_{ijt} = 1$).

3.3 Control function

To correct for this bias, we develop a control function approach (Wooldridge 2015). We begin by noting the conditional distribution of $(\varepsilon_{i0t} - \varepsilon_{ijt})$ in the selected sample of reservers. Let $f(\cdot)$ be the logistic density and $F(\cdot)$ the logistic distribution and define $\tilde{\varepsilon}_{ijt} \equiv (\varepsilon_{i0t} - \varepsilon_{ijt})$.

The conditional density of $\tilde{\varepsilon}_{ijt}$ is

$$
\begin{aligned}
f(\ \tilde{\varepsilon}_{ijt} \mid R_{ijt} = 1) &= f(\ \tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt} - \bar{V}_{i0t}) \\
&= \frac{f(\tilde{\varepsilon}_{ijt}) \cdot \mathbb{1}\{\tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt} - \bar{V}_{i0t}\}}{F(\bar{V}_{ijt} - \bar{V}_{i0t})} \\
&= \frac{f(\tilde{\varepsilon}_{ijt}) \cdot \mathbb{1}\{\tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt}\}}{\mathbb{P}(R_{ijt} = 1)}, \qquad (9)
\end{aligned}
$$

where the first line follows from the reservation condition in Equation 5, the second from the definition of a truncated density, and the third by noting that $V_{i0t} = 0$ and that $F(\bar{V}_{ijt}) = \mathbb{P}(R_{ijt} = 1)$.

An estimand for $\tilde{\varepsilon}_{ijt}$ is given by

$$
\begin{aligned}
\mathbb{E}[\ \tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt}] &= \int_{-\infty}^{\infty} \tilde{\varepsilon}_{ijt}\ f(\tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt}])\ d\tilde{\varepsilon}_{ijt} \\
&= \frac{\int_{-\infty}^{\bar{V}_{ijt}} \tilde{\varepsilon}_{ijt}\ f(\tilde{\varepsilon}_{ijt})\ d\tilde{\varepsilon}_{ijt}}{\mathbb{P}(R_{ijt} = 1)} \\
&= \frac{\bar{V}_{ijt} \cdot \frac{\exp(\bar{V}_{ijt})}{1+\exp(\bar{V}_{ijt})} - \log(1 + \exp(\bar{V}_{ijt}))}{\mathbb{P}(R_{ijt} = 1)} \\
&= \frac{\bar{V}_{ijt} \cdot \mathbb{P}(R_{ijt} = 1) - I_{ijt}}{\mathbb{P}(R_{ijt} = 1)} \\
&= \bar{V}_{ijt} - \frac{I_{ijt}}{\mathbb{P}(R_{ijt} = 1)}. \qquad (10)
\end{aligned}
$$

The first line follows from the definition of a conditional expectation, the second by substituting in Equation 9, the third by evaluating the definite integral, the fourth by substituting Equation 5 and defining $I_{ijt} \equiv \log(1 + \exp(\bar{V}_{ijt}))$, and the last through simplification.

Equation 10 contains familiar terms. The $\bar{V}_{ijt}$ term gives the expected utility of the site choice from the reservation decision. The second term contains the inclusive value I_{ijt}, which is equivalent to the expected maximal utility a visitor could expect from holding the reservation, including the value of either the trip or the cancellation (Train 2009). The I_{ijt}

term is scaled by the inverse of the probability that they would reserve at the site; higher reservation probabilities result in higher estimates of preferences $\tilde{\varepsilon}_{ijt}$.

Including $\tilde{\varepsilon}_{ijt}$ as a control function within Equation 8 provides an estimate of the unobserved preferences of individual i and allows for unbiased estimation of the travel cost parameter in the cancellation problem. As travel cost is positively correlated with ε_{ijt}, we expect that it is negatively correlated with $\tilde{\varepsilon}_{ijt} \equiv (\varepsilon_{i0t} - \varepsilon_{ijt})$. We also expect a higher value of $\tilde{\varepsilon}_{ijt}$ to increase the likelihood of cancellation, as in Equation 8. In Section 3.4, we illustrate the bias correction of this control function through a numerical example.

Estimation of the cancellation decision proceeds through the following two-stage process. First, we estimate the parameters of the reservation decision $\mathbb{P}(R_{ijt} = 1)$ by maximizing a zonal log-likelihood function as in Equation 6, for reservations made earlier than $t - \tau$ and using expected site conditions. Then, we use the parameters to create a fitted value $\hat{\tilde{\varepsilon}}_{ijt}$ for every observed reservation. We substitute them into the trip-level equation for the cancellation decision, where each row of data is a trip with a dependent variable $C_{ijt} \in \{0, 1\}$ indicating whether the user cancelled. In this second stage, the independent variables in V_{ijt} use realized rather than expected site conditions, as users approximately know the site conditions close to the arrival date. For individual i, the log-likelihood function for the cancellation decision is

$$\ell^C(\omega, \rho) = \sum_{i=1}^{N} \sum_{j=0}^{J} \sum_{t=1}^{T} (1 - C_{ijt}) \log \big(\mathbb{P}(C_{ijt} = 0 | \omega, \rho, R_{ijt} = 1)\big) + \\ C_{ijt} \log \big(1 - \mathbb{P}(C_{ijt} = 0 | \omega, \rho, R_{ijt} = 1)\big). \quad (11)$$

Because of the two-stage estimation, we bootstrap the main estimates to obtain appropriate standard errors (Cameron and Miller 2015; Wooldridge 2015).

3.4 Numerical example

To illustrate the source of bias in naïve cancellation estimates and the ability of our control function estimator to correct for this bias, we test our estimator on simulated data. We simulate 10,000 draws with $N = 100,000$ users (individuals i) who make sequential reservation and cancellation decisions. We assign each user random travel costs, smoke conditions, and unobserved preferences ε_{ijt} and υ_{ijt}. We assert an arbitrary true WTP to avoid smoke of $\phi/\delta = 2$. Appendix D provides additional details.

Table 1 summarizes results from the simulation and illustrates two dimensions of the identification challenge. First, we test the role of sample selection, using the fact that in simulated data, we can observe the counterfactual cancellation decisions of users who never held a reservation. In Column 1, we estimate WTP based on the full sample of reservers and nonreservers. With no sample selection (the sample is not limited to reservers), our estimate approaches 2, the true WTP. In Column 2, the sample is limited to reservers, but we model cancellation and reservation decisions as independent, conditional on observables. That is, unobserved preferences affecting reservation decisions do not affect cancellation decisions, and second-stage unobserved preferences are equal to $\frac{1}{\rho}\eta$ rather than $\varepsilon_{it} + \frac{1}{\rho}\eta$. Under this assumption, our estimate is again nearly equal to the true WTP. It is only in Column 3, when the sample is selected and preferences in the first and second stages are not conditionally independent, that WTP is biased. In Appendix D, we discuss how this bias operates through correlation between preferences and travel cost in the selected sample, which attenuates estimates of the travel cost parameter. In Column 4, we maintain both assumptions from Column 3 but also introduce our control function for $\tilde{\varepsilon}_{ijt}$. Across Monte Carlo simulations, the control function corrects the bias and includes the true WTP in the confidence interval. For a full treatment, refer to Appendix D.

Table 1: Numerical Example for 10,000 Simulations of Cancellation Estimation, Bias, and Bias Correction from $\tilde{\varepsilon}_{ijt}$ Control Function

	(1)	(2)	(3)	(4)
WTP	2.01**	2.01**	6.70	2.00**
	(0.14)	(0.15)	(8.52)	(0.25)
Sample	All users	Reservers	Reservers	Reservers
Second-stage errors	$\varepsilon_{ijt} + \frac{1}{\rho}\eta_{ijt}$	$\frac{1}{\rho}\eta_{ijt}$	$\varepsilon_{ijt} + \frac{1}{\rho}\eta_{ijt}$	$\varepsilon_{ijt} + \frac{1}{\rho}\eta_{ijt}$
Control function	No	No	No	Yes

Notes: True willingness to pay (WTP) = 2. WTP values are means (with standard errors in parentheses) across estimates from M = 10,000 simulations with N = 100,000. * $p < 0.05$, ** $p < 0.01$.

4 Estimation

In this section, we estimate the welfare damages of wildfire smoke for outdoor recreation. As discussed in the previous section, the estimation follows a two-stage process that links reservations to cancellations close to arrival. Figure 3 shows the timing of decisions. We restrict the data to the set of users who booked more than a week ahead of time, or $\tau = 7$ in Figure 3, and decided whether to cancel within a week of the arrival date. We therefore exclude reservations that were cancelled more than a week in advance. We also focus on trips scheduled for the months of May–September and over the years 2010–2017. Last, we limit attention to trips coming from within 650 km (400 miles), as described in Section 2.6. These restrictions result in a sample of 2,723,940 reservations.[20]

4.1 Cancellations close to arrival

Figure 4 displays how the cancellation rate varies by travel cost and wildfire smoke conditions. The figure shows that users cancel their trips at higher rates during smoke conditions. This relationship does not appear to vary by travel cost, as the distance between the red and blue points is relatively constant across travel cost bins. Visually, the slope between cancellation rate and travel cost appears shallow. As explored in Section 3 and Appendix D, this shallow slope is likely due to positive correlation between travel cost and the unobserved preference parameter ε_{ijt} among the selected sample of reservers. Intuitively, if we were to observe someone reserve at site j despite a high travel cost, on average, they should have a higher preference ε_{ijt} for the site than someone with a similar travel cost who did not reserve, such that $\mathbb{E}[\varepsilon_{ijt}c_{ijt}|R_{ijt} = 1] > 0$. If ignored, we expect this correlation to depress the magnitude of the travel cost coefficient in the estimation of cancellations $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$, which translates to a shallow slope in Figure 4.

[20] A "reservation" or "trip" is composed of multiple "transactions," which could include an initial booking, payment, check-in, cancellation, or refund.

Figure 4: Cancellation Rate Close to Arrival

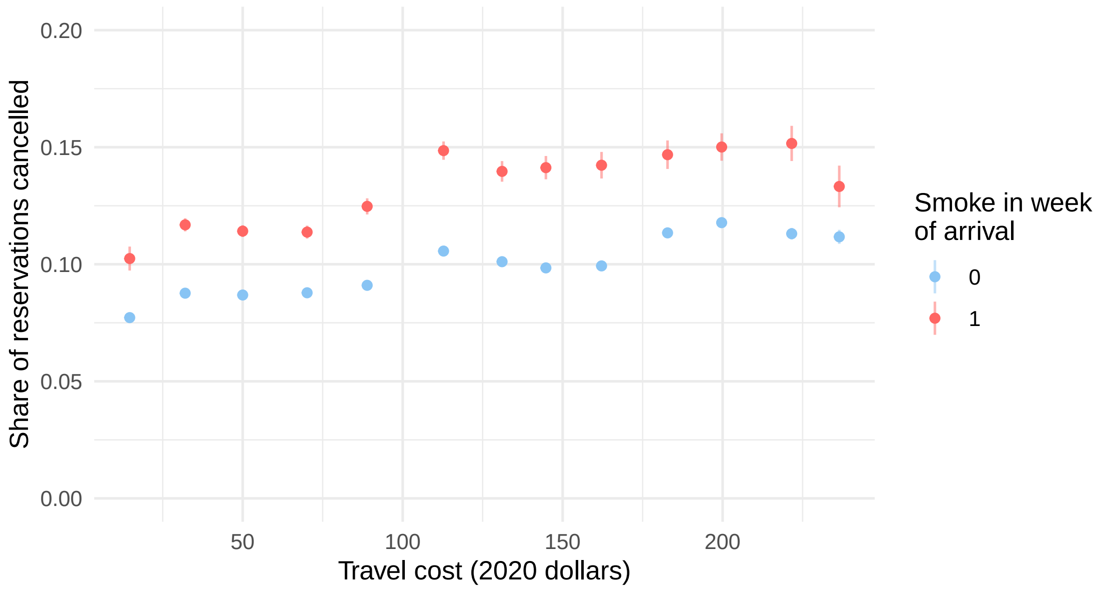


Table 2 reports results for biased estimation of cancellations $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ using the trip-level maximum likelihood function of Equation 11. These estimates ignore the correlation between ε_{ijt} and travel cost among the set of users who chose to reserve. WTP is computed by taking the ratio between marginal disutility of smoke to that of expenditure (the smoke coefficient divided by the travel cost coefficient). Standard errors for WTP are computed using the delta method. In all estimations, the observations are weighted via frequency weights because a single reservation might represent, for example, two or eight visitors.

In Column 1, we display results without controlling for campground or seasonal fixed effects. Columns 2–4 add fixed effects. We include a campground fixed effect to account for location-specific, time-invariant unobservables related to site quality. We also account for differences in reservation rates based on the day of the week, as weekends see higher reservation activity. A campground-by-week fixed effect controls for unobserved location-specific seasonality, such as seasonal natural phenomena. Last, we include various year fixed effects to account for time-related unobservables. Column 4 would imply that wildfire smoke causes \$154 in lost welfare per person per trip. This result is likely upward biased because WTP $= \phi/\delta$, and we expect the travel cost parameter δ to be attenuated.

Table 2: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within One Week, Uncorrected for Sample Selection

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2195**	-0.2615**	-0.2346**	-0.2615**
	(0.0238)	(0.0283)	(0.0273)	(0.0215)
Travel cost (dollars)	-0.0024**	-0.0017**	-0.0017**	-0.0017**
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Inv. distance to wildfire (km^{-1})	-11.1276**	-12.0389**	-11.9174**	-7.8003**
	(0.9266)	(2.4288)	(2.4432)	(0.8291)
High temp. (degrees C)	0.0198**	0.0287**	0.0292**	0.0307**
	(0.0045)	(0.0023)	(0.0023)	(0.0022)
Low temp. (degrees C)	-0.0033	-0.0205**	-0.0214**	-0.0253**
	(0.0058)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0041**	-0.0058**	-0.0060**	-0.0057**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
N	2,723,830	2,692,468	2,692,468	2,689,216
WTP	91.1**	153.4**	137.35**	154.04**
	(12.36)	(21.06)	(19.85)	(15.43)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

To correct for the biased WTP in Table 2, we use the control function described in Equation 10, $\tilde{\varepsilon}_{ijt} = \bar{V}_{ijt} - \frac{I_{ijt}}{\mathbb{P}(R_{ijt}=1)}$. The first step is to estimate the probability of reservation earlier than one week based on expected site conditions and using a zonal travel cost model. Then, we fit the parameters from the reservation estimation to form an estimate for $\tilde{\varepsilon}_{ijt}$. This estimate is used as a covariate in the trip-level estimation of cancellations $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, after site conditions become approximately known to visitors.

We construct expected site conditions in the following way. For temperature and precipitation, we use climate normals from our PRISM data source, which represent average weather conditions for 1980–2010. Because travel cost is likely known to the individual ahead of time, we use the visitor's actual travel cost. For expected smoke and expected distance to fire, we use the average conditions over the past four years. For example, if a site was affected by smoke for one out of the past four years, we code expected smoke as 0.25.

Table 3 shows results from the first-stage reservation decision $\mathbb{P}(R_{ijt} = 1)$ implied by Equation 6. Users appear unexpectedly more likely to reserve at a campground with a higher expectation of wildfire smoke. Including more fixed effects generally decreases the magnitude and significance of the estimate, including moving the WTP closer to 0. Still, even with a high number of seasonal fixed effects, we may be unable to remove the correlation of seasonal variation in camping with wildfire smoke. Nevertheless, the primary purpose to estimate the likelihood of reservation $\mathbb{P}(R_{ijt} = 1)$ is as an input for the control function $\tilde{\varepsilon}_{ijt}$ in the estimation of $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$, so we should be unconcerned by the direction of the smoke expectation parameter.

Table 3: $\mathbb{P}(R_{ijt} = 1)$ for Reservations Made Earlier Than One Week Based on Expected Site Conditions

	(1)	(2)	(3)	(4)
Smoke exp.	0.9260**	0.2513**	0.1032**	0.0822*
	(0.0036)	(0.0423)	(0.0363)	(0.0324)
Travel cost (dollars)	-0.0202**	-0.0244**	-0.0244**	-0.0244**
	(0.0000)	(0.0013)	(0.0013)	(0.0013)
Inv. distance to wildfire exp. (km^{-1})	39.6901**	6.0569**	6.1590**	6.7856**
	(0.0742)	(1.7451)	(1.5707)	(1.4654)
High temp. exp. (degrees C)	0.0191**	0.0597**	0.0611**	0.0588**
	(0.0001)	(0.0130)	(0.0131)	(0.0126)
Low temp. exp. (degrees C)	-0.0191**	-0.0818**	-0.0835**	-0.0812**
	(0.0001)	(0.0153)	(0.0153)	(0.0148)
Precip. exp. in week of arrival (mm)	-0.0126**	0.0071**	0.0066*	0.0067*
	(0.0001)	(0.0027)	(0.0027)	(0.0027)
N	15,209,187	12,668,366	12,668,366	12,298,572
WTP	-45.93**	-10.31**	-4.23**	-3.37*
	(0.18)	(1.72)	(1.45)	(1.31)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

After zonal estimation of $\mathbb{P}(R_{ijt} = 1)$ for early reservers, we use the parameter estimates to create fitted probabilities of reservation at the trip level. Figures A8 and A9 in Appendix A show the variation in fitted probability $\mathbb{P}(R_{ijt} = 1)$ and in the control function $\tilde{\varepsilon}_{ijt}$. We expect that preferences and travel costs are correlated in the selected sample, $\mathbb{E}[c_{ijt}\varepsilon_{ijt}|R_{ijt} = 1] > 0$, so our control function should be inversely correlated with travel cost, $\mathbb{E}[c_{ijt}(\varepsilon_{i0t} - \varepsilon_{ijt})|R_{ijt} = 1] < 0$. Figure A10 in Appendix A illustrates this correlation empirically using the fitted values of $\tilde{\varepsilon}_{ijt}$ and the travel cost for the sample of reservers. This empirical result is consistent with the prediction of our theory and numerical exercise in Section 3 and Appendix D.

Table 4 reports the main trip-level results for the cancellation estimation $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ using the bias correcting control function $\tilde{\varepsilon}_{ijt}$. The coefficient for $\tilde{\varepsilon}_{ijt}$ is significant, suggesting that preferences at the time of reservation are an important determinant of the cancellation decision. In addition, comparing to Table 2, the travel cost coefficient was the only parameter to change when including $\tilde{\varepsilon}_{ijt}$, which is consistent with sample selection bias operating through correlation with travel cost. Overall, the WTP estimates are reduced to \$107 per person per trip of lost utility due to cancellations. By comparison, the biased results in Table 2 were \$154 per person per trip, which is 44 percent higher.

Wooldridge (2015) recommends bootstrapping standard errors for control functions because of the two-stage estimation process. For the main estimates in Table 4, we follow the clustered bootstrapping process of Cameron and Miller (2015), drawing with replacement at the campground level for 400 bootstraps. In Appendix E, we report results from Shapiro-Wilk tests for normality, failing to reject the null hypothesis that the bootstrapped smoke coefficients and travel cost coefficients are normally distributed. These tests suggest that 400 bootstraps are adequate for the analysis.

4.2 Relationship of damages to smoke duration

In this section, we investigate how estimated welfare losses vary according to one measure of the severity of smoke impacts: the number of smoke days in the week of arrival. We hypothesize that welfare impacts could vary with this measure for two reasons. First, it is likely that sites experiencing a greater number of smoke days in close succession also experience a larger degradation in air quality. Second, additional smoke days in the week of

Table 4: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within One Week, Corrected for Sample Selection

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2175**	-0.2708**	-0.2438**	-0.2603**
	(0.0247)	(0.0238)	(0.0221)	(0.0218)
Travel cost (dollars)	-0.0026**	-0.0024**	-0.0024**	-0.0025**
	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-11.1017**	-10.8883**	-10.7067**	-7.8141**
	(0.8580)	(1.4280)	(1.4288)	(0.7920)
High temp. (degrees C)	0.0202**	0.0284**	0.0289**	0.0306**
	(0.0043)	(0.0024)	(0.0024)	(0.0023)
Low temp. (degrees C)	-0.0037	-0.0204**	-0.0214**	-0.0252**
	(0.0052)	(0.0026)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0041**	-0.0058**	-0.0060**	-0.0057**
	(0.0010)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0112	-0.0356**	-0.0366**	-0.0385**
	(0.0284)	(0.0106)	(0.0105)	(0.0106)
N	2,723,034	2,691,655	2,691,655	2,688,739
WTP	85.23**	113.91**	101.50**	107.14**
	(17.82)	(18.48)	(16.50)	(16.33)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Bootstrapped std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

arrival communicate a higher likelihood of smoke on the actual arrival date.[21]

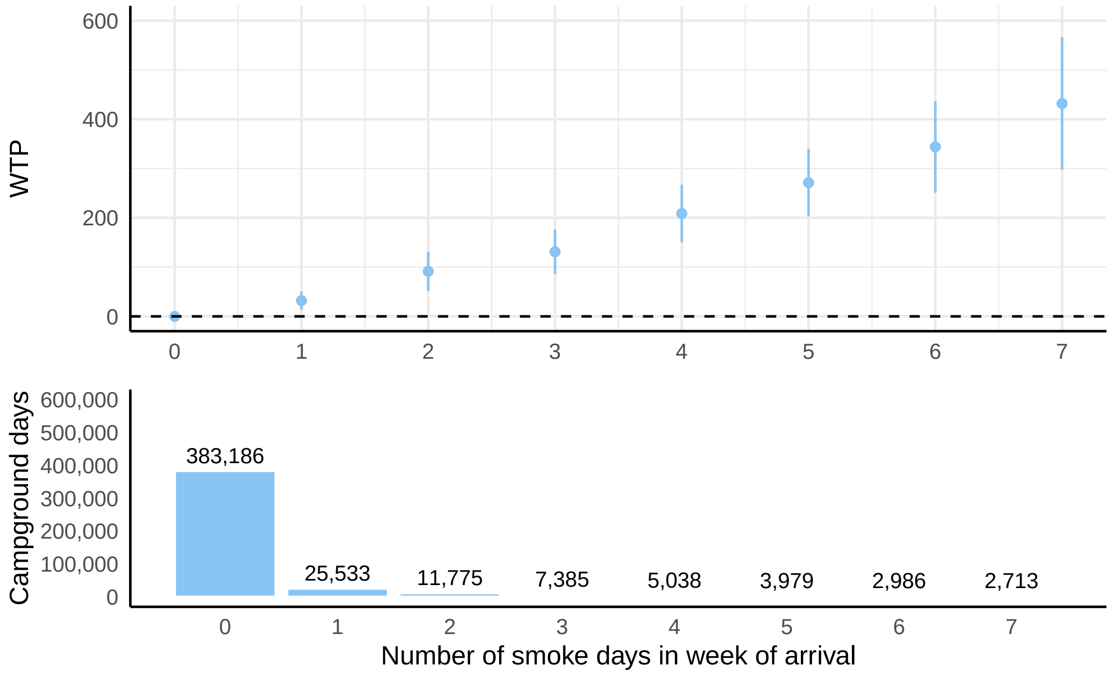
To investigate heterogeneity in estimated welfare impacts, we respecify $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ to allow for differential effects based on the number of smoke-affected days in the week of arrival. Figure 5 plots the resulting WTP estimates.[22] Damages monotonically increase in the number of smoke days, giving confidence that the estimates reflect damages from smoke. When a campground was affected by smoke on all seven days in the week of arrival, we find welfare damages of \$432 per person per trip. Estimated welfare impacts are likely

[21] Table A2 in Appendix A reports results from a regression of $\mathbb{1}\{\text{campground is smoke-affected}\}$ on indicator variables for one, two, ..., seven days of smoke in the week before arrival. Each additional smoke day increases the probability of smoke on the actual day of arrival. For instance, compared to a week with no smoke days, a week with two days raises the probability by 0.301; a week with six days raises it by 0.739.

[22] Full results are included in Table A3 in Appendix A.

greater during more smoke-affected weeks due to worse air quality conditions or because the perceived likelihood of smoke during the visit is greater.

Figure 5: Greater Welfare Damages for Weeks That Were More Smoke Affected, Consistent with Either More Severe Events or Increased Certainty of Smoke Conditions



4.3 Placebo test for smoke

As a robustness check, we devise a placebo test to check whether the smoke coefficient actually measures responses to smoke. The placebo considers the responses of visitors whose campground was not affected by smoke until one or two weeks after their arrival. If visitors are truly averting recreation due to smoke, then we should see no response to these placebos. Of the 2.38 million reservations without smoke in the week of arrival, more than 375,000 placebo reservations are created.

Table 5 displays results from the placebo test. Across the main specifications, we find null responses to the two smoke placebos. Comparing to Table 4, most coefficients remain the same for this placebo test. This exercise should add confidence that individuals are actually responding to smoke in the main estimation.

Table 5: Placebo Test for $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ Using Smoke Long After Arrival

	(1)	(2)	(3)	(4)
Smoke in week after arrival	0.0874**	0.0248	0.0226	0.0066
	(0.0190)	(0.0163)	(0.0155)	(0.0159)
Smoke two weeks after arrival	0.0783**	0.0041	0.0034	-0.0042
	(0.0231)	(0.0147)	(0.0146)	(0.0156)
Travel cost (dollars)	-0.0027**	-0.0025**	-0.0025**	-0.0025**
	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-8.5921**	-7.3229**	-7.2413**	-5.2047**
	(0.8373)	(0.8998)	(0.8823)	(0.7795)
High temp. (degrees C)	0.0201**	0.0294**	0.0285**	0.0303**
	(0.0043)	(0.0022)	(0.0022)	(0.0022)
Low temp. (degrees C)	-0.0010	-0.0184**	-0.0184**	-0.0221**
	(0.0057)	(0.0026)	(0.0026)	(0.0025)
Precip. in week of arrival (mm)	-0.0039**	-0.0062**	-0.0062**	-0.0058**
	(0.0011)	(0.0009)	(0.0009)	(0.0010)
$\tilde{\varepsilon}_{ijt}$	-0.0089	-0.0332**	-0.0337**	-0.0352**
	(0.0262)	(0.0123)	(0.0122)	(0.0124)
N	2,379,842	2,344,620	2,344,620	2,340,894
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

4.4 Additional results and robustness checks

We report several additional results and robustness checks in the appendices. The first analysis explores how results differ when excluding observations with nearby active wildfires. In addition, we assess the role of no-shows in the estimation of cancellation probability. A third exercise varies the distance threshold that defines the sample restriction. We also vary the temporal threshold to consider cancellations. Last, we show how results vary by the popularity of the recreation destination.

The literature has found that visitors to national parks are less avoidant of wildfire smoke that originates from distant sources (Cai 2021). In Appendix F, we investigate how nearby active wildfires affect the main estimates. Although the main estimation controls for proximity to active wildfire, one might still be concerned that individuals avoid recreation due to fire rather than smoke. If smoke days are highly correlated with nearness to fire, it could increase the estimated smoke coefficient and inflate WTP. To address this possibility, we reestimate the main specifications but remove observations with a nearby active fire. When removing observations with a fire within 20 km (12 miles), we find reduced welfare damages of \$85. These results are consistent with the broader findings of Cai (2021).

Another concern when studying cancellations is whether an individual will formally cancel or simply not show up. For most campgrounds, we do not observe whether an individual checks in. However, campers have an incentive to cancel their reservation. For cancellations made more than 24 hours before the arrival date, visitors are reimbursed for the full cost less a \$10 cancellation fee; when cancelling within 24 hours of arrival, they are reimbursed for the full trip less the \$10 fee and the price of the first night's stay. Still, we explore this question in Appendix G. For a small subset of campgrounds, we are able to observe no-shows; in this sample, we demonstrate that including or excluding no-show observations in the estimation of cancellations $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ does not change the estimates for the smoke or travel cost coefficients. For a discussion of this issue, see Appendix G.

We also explore alternative distance thresholds for the sample restriction. In the main results, we limit attention to reservations made within 650 km of one-way driving distance, or approximately 400 miles. Figure 1 shows that this distance restriction includes more than 85 percent of all reservations. Appendix H reports how estimates vary with this threshold. Increasing the distance threshold attenuates the parameter estimate for travel cost, which is an input to welfare calculation. This may be due to the inclusion of visitors traveling from greater distances, some of whom were on multipurpose trips and were therefore less likely to cancel. As a result, the estimated welfare damages increase as the distance threshold is relaxed. For more information, see Appendix H.

In addition, we vary the threshold value for τ in Figure 3. The main estimation considers cancellation decisions within $\tau = 7$ days of arrival, but Appendix I reports results for alternative thresholds of 3, 5, and 9 days. As in the main estimation, the variable of interest is a binary variable equal to 1 for a smoke-affected day within τ days of arrival. Welfare estimates are larger for shorter bandwidths of τ. For $\tau = 3$, we find estimates of

$137 per person per trip; for $\tau = 9$, we find estimates of $92 per person per trip. These results point to a similar mechanism as in Section 4.2. During shorter time windows, the occurrence of smoke communicates a greater likelihood of smoke on the actual arrival date. An information effect is consistent with greater smoke avoidance closer to the arrival date.

Last, we assess which types of campgrounds drive the parameter estimates. Even with a high number of fixed effects, visitors could respond differentially to disamenities at highly valued destinations, such as Glacier National Park, versus small, primarily local US Forest Service campgrounds. Appendix J reports heterogeneous results by campground popularity, where popularity is defined by annual visitation. We find that visitors are less responsive to both smoke and travel cost at the most popular destinations. Visitors may be more tolerant of environmental disamenities at highly valued destinations. Across most specifications, welfare damages are highest for destinations in the middle quartiles of popularity.

5 Total welfare losses

In the preceding sections, we estimated per trip damages of wildfire smoke. We now turn to an appraisal of the total annual welfare damages for recreation. We combine the camping data from Recreation.gov with overall visitation data from federal and state agencies to determine the total number of outdoor visits in the west that are affected by smoke each year. As a back-of-the-envelope calculation, we multiply total smoke-affected visitation by the empirical per trip welfare estimate to approximate the total annual welfare loss due to smoke in the West. One limitation of this analysis is that the estimates are derived from camping activity, which may not be representative of losses to other forms of recreation, such as angling, swimming, or daytime visits. Still, this figure approximates the relative magnitude of total annual smoke damages for recreation in the western United States.

We find that across federal and state lands, an average of 21.5 million outdoor recreation visits per year are affected by wildfire smoke. Multiplying by a per trip damage of $107 per person, this result implies more than $2.3 billion of welfare losses each year. This back-of-the-envelope estimate represents the lost welfare to inframarginal visitors and does not include the value of lost trips.

To arrive at this number, we use total visitation numbers from the National Park Service,[23] US Forest Service,[24] Bureau of Land Management,[25] US Army Corps of Engineers,[26] and National Association of State Park Directors (Smith et al. 2019) for 2008–2017. These data sources have varying levels of spatial and temporal granularity. For each data source, we use the Recreation.gov data to determine, at the relevant spatial and temporal scale, the proportion of total visits at each agency that were affected by smoke. For more information on the estimation of smoke-affected visitation, see Appendix K.

Table 6 displays estimates of total visitation, smoke-affected visitation, and total welfare losses by agency. One key point is the high overall level of outdoor recreation, with more than 511 million annual visits to state and federal lands in the western United States. In addition, a high proportion of these visits are affected by smoke. We estimate that approximately 21.5 million visits per year are smoke affected, or 4.2 percent. When multiplied by the per-trip estimate of $107, we find total annual welfare losses of approximately $2.3 billion. Nearly half of these damages occur at state parks, which see larger visitation compared to federal agencies. Of any agency, the US Army Corps of Engineers saw the highest proportion of its visitors affected by smoke. This is likely because much of that agency's visitation (nearly 40 percent) occurs at lakes and reservoirs in the Pacific Northwest, a region that has seen particularly high wildfire smoke impacts relative to other regions (Burke et al. 2021; Gellman et al. 2022; Miller et al. 2021).

Welfare losses vary by region. Some states saw high smoke damages due to high baseline levels of visitation, and damages in other regions were driven by a high proportion of smoke-affected visits. Panel B of Table 6 reports losses by state. For states such as California and Colorado, damages are large due to high visitation. States such as Oregon and Washington saw both relatively high visitation and a high share of smoke-impacted visits. At the other end of the spectrum, states in the Southwest, such as Arizona, Nevada, and Utah, saw high visitation but a low proportion of smoke-affected visits. In Northern Rocky Mountain states,

[23] National Park Service. Annual Summary Report. `https://irma.nps.gov/STATS`.

[24] US Forest Service. National Visitor Use Monitoring Program. `https://www.fs.usda.gov/about-agency/nvum`.

[25] Bureau of Land Management. Public Land Statistics. `https://www.blm.gov/about/data/public-land-statistics`.

[26] US Army Corps of Engineers. Value to the Nation. `https://www.iwr.usace.army.mil/Missions/Value-to-the-Nation`.

such as Idaho, Montana, and Wyoming, damages are driven by a high share of smoke-affected days despite lower total visitation. Figure A11 in Appendix A maps the results to contrast proportional versus total impacts. As a whole, these findings show the high cost of wildfire smoke for outdoor recreation in the western United States.

Table 6: Smoke-Affected Recreation Visits and Welfare Losses for the Western United States, 2008–2017

	Total visits/year (millions)	Smoke-affected visits/year (millions)	Proportion visits smoke-affected	Welfare loss/year (millions)
Panel A: Agency				
National Park Service	102.6	2.3	0.023	\$248.1
US Forest Service	108.0	4.8	0.044	\$511.4
Bureau of Land Management	59.8	2.5	0.042	\$267.3
US Army Corps of Engineers	46.4	2.4	0.051	\$251.7
State Parks	194.5	9.6	0.049	\$1,022.1
Total	511.4	21.5	0.042	\$2,300.6
Panel B: State				
Arizona	33.0	0.4	0.013	\$46.3
California	162.7	6.1	0.037	\$649.6
Colorado	55.3	2.1	0.037	\$220.9
Idaho	19.6	1.3	0.065	\$136.8
Montana	18.3	1.1	0.062	\$121.8
New Mexico	13.7	0.6	0.041	\$60.9
Nevada	20.2	0.3	0.014	\$29.8
Oregon	69.9	4.4	0.062	\$466.2
Utah	32.1	0.8	0.025	\$84.8
Washington	64.8	3.3	0.051	\$351.9
Wyoming	21.9	1.2	0.056	\$131.6
Total	511.4	21.5	0.042	\$2,300.6

Note: Welfare losses computed by multiplying \$107 per trip by smoke-affected visits.

6 Conclusion

This study provides the first revealed preference welfare estimates of the damage of wildfire smoke for outdoor recreation. Using high-frequency data on campground reservations, wildfire, smoke, and air pollution, we study avoidance behavior at federally managed lands in the western United States. We estimate that wildfire smoke causes welfare losses of $107 per person per trip. These damages increase at an increasing rate when campgrounds are affected by consecutive days of smoke and are attenuated when smoke-affected campgrounds are far from active wildfires. Combining these results with federal and state data on total visitation, we estimate that 21.5 million outdoor recreation visits per year are affected by smoke, with associated welfare losses of $2.3 billion.

The paper provides several contributions to the literature. First, we use novel methods and data. We value a temporary environmental bad, wildfire smoke, in a context where visitors face changing sets of information. We develop a two-stage decision structure that links preferences with a control function. This model draws on work from economists concerned with sample selection in nonlinear models (Greene 2012; Terza 2009) and researchers confronting sample selection in recreation settings (Cameron and DeShazo 2013; Cameron and Kolstoe 2022; Kolstoe and Cameron 2017; Lewis et al. 2019). The framework we develop could be used in other studies facing sample selection or sequential choices. Furthermore, our use of administrative data complements recent literature using large or innovative datasets to study recreation across multistate regions (Cameron and Kolstoe 2022; Dundas and von Haefen 2020; English et al. 2018; Parthum and Christensen 2022).

We also add to the literature on the costs of wildfire smoke. To contextualize the results of this study, we compare to several other studies on the costs of wildfire smoke. Most of these studies used survey methods or health care costs or valued changes in mortality using VSL. Richardson et al. (2012) report results from a survey following a large wildfire in Los Angeles County. They asked respondents about avoidance behavior during this fire (expenditures on air purifiers) and health outcomes and risk perceptions. They derive WTP to avert one wildfire-induced symptom day of $84 in 2009 dollars. We estimate WTP to avoid an exposure day rather than a symptom day. Taking the empirical estimate of $107 per trip, this translates to approximately \$38 per day based on an average trip of 2.84 days.

We can also compare total welfare results to the literature. We estimate welfare losses of approximately \$2.3 billion per year for recreation in the western United States. Miller et al. (2021) combined a VSL with estimates of mortality among elderly Medicare recipients due to wildfire smoke. They found \$6–170 billion in annual damages, in 2021 dollars. These results mainly vary due to assumptions on remaining years of life, as their sample is comprised largely of elderly individuals. When assuming that those who die from wildfire smoke would have lived an additional 3.5 years, they arrive at a lower bound of \$6 billion. Borgschulte et al. (2022) found annual lost labor earnings of \$125 billion per year, in 2018 dollars, due to wildfire smoke. Several other studies have found costs of wildfire smoke for test scores, crime, and hospital visits (Burkhardt et al. 2019; Cullen 2020; Wen and Burke 2022).

Estimating these costs can inform public policy. The federal government spends an average of \$2.8 billion per year on fire suppression, and the State of California spends \$900 million per year.[27,28] Wildfires destroy thousands of structures per year, which has cost tens of billions of dollars in recent years (Baylis and Boomhower 2022, 2023; Buechi et al. 2021). Both states and the federal government have pledged to increase fuel treatment projects to mitigate the risk of fire ignition and spread. California has jointly declared a goal with the US Forest Service to treat more than 1 million acres of hazardous vegetation per year[29] and proposed to spend \$1.2 billion across fiscal years 2022–23 and 2023–24 for fire mitigation activities, such as vegetation management and home hardening.[30] Understanding the cost of wildfires is crucial to assess the benefit of these public policies. Our study contributes to a growing understanding of the costs of wildfire smoke.

---

[27] National Interagency Fire Center. Suppression Costs. `https://www.nifc.gov/fire-information/statistics/suppression-costs`.

[28] California Department of Forestry and Fire Protection. Suppression Costs. `https://www.fire.ca.gov/stats-events`.

[29] Agreement for shared stewardship of California's forest and rangelands between the State of California and the USDA Forest Service, Pacific Southwest Region. `https://www.gov.ca.gov/wp-content/uploads/2020/08/8.12.20-CA-Shared-Stewardship-MOU.pdf`.

[30] California Legislative Analyst's Office. The 2022–23 Budget Wildfire and Forest Resilience Package. `https://lao.ca.gov/Publications/Report/4495`.

7 References

Abatzoglou, J.T., and A.P. Williams. 2016. Impact of Anthropogenic Climate Change on Wildfire Across Western US Forests. *Proceedings of the National Academy of Sciences* 113(42): 11770–75.

Aguilera, R., T. Corringham, A. Gershunov, and T. Benmarhnia. 2021. Wildfire Smoke Impacts Respiratory Health More Than Fine Particles from Other Sources: Observational Evidence from Southern California. *Nature Communications* 12(1): 1–8.

Baylis, P., and J. Boomhower. 2022. *Mandated vs. Voluntary Adaptation to Natural Disasters: The Case of US Wildfires* (No. w29621). National Bureau of Economic Research.

Baylis, P., and J. Boomhower. 2023. The Economic Incidence of Wildfire Suppression in the United States. *American Economic Journal: Applied Economics* 15(1): 442–73.

Berry, S.T. 1994. Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics* 25(2): 242–62.

Borgschulte, M., D. Molitor, and E.Y. Zou. 2022. Air Pollution and the Labor Market: Evidence from Wildfire Smoke. *Review of Economics and Statistics*: 1–46.

Buechi, H., P. Weber, S. Heard, D. Cameron, and A.J. Plantinga. 2021. Long-Term Trends in Wildfire Damages in California. *International Journal of Wildland Fire* 30(10): 757–62.

Burke, M., A. Driscoll, S. Heft-Neal, J. Xue, J. Burney, and M. Wara. 2021. The Changing Risk and Burden of Wildfire in the United States. *Proceedings of the National Academy of Sciences*: 118(2).

Burke, M., S. Heft-Neal, J. Li, A. Driscoll, P.W Baylis, M. Stigler, J. Weill, and C. Gould. 2022. Exposures and Behavioral Responses to Wildfire Smoke. *Nature Human Behaviour* 6(10): 1351–61.

Burkhardt, J., J. Bayham, A. Wilson, E. Carter, J.D. Berman, K. O'Dell, B. Ford, E.V. Fischer, and J.R. Pierce. 2019. The Effect of Pollution on Crime: Evidence from Data on Particulate Matter and Ozone. *Journal of Environmental Economics and Management* 98: 102267.

Cai, C. 2021. Wildfire and Visitation in US National Parks. Working paper, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign.

Cameron, A.C., and D.L. Miller. 2015. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources* 50(2): 317–72.

Cameron, T.A., and J.R. DeShazo. 2013. Demand for Health Risk Reductions. *Journal of*

Environmental Economics and Management 65(1): 87–109.

Cameron, T.A., and S.H. Kolstoe. 2022. Using Auxiliary Population Samples for Sample-Selection Correction in Models Based on Crowd-Sourced Volunteered Geographic Information. *Land Economics* 98(1): 1–21.

Cullen, D. 2020. The Health Cost of Wildfire Smoke. Working paper, Department of Economics, University of California, Santa Barbara.

Dundas, S.J., and R.H. von Haefen. 2020. The Effects of Weather on Recreational Fishing Demand and Adaptation: Implications for a Changing Climate. *Journal of the Association of Environmental and Resource Economists* 7(2): 209–42.

English, E., R.H. von Haefen, J. Herriges, C. Leggett, F. Lupi, K. McConnell, M. Welsh, A. Domanski, and N. Meade. 2018. Estimating the Value of Lost Recreation Days from the Deepwater Horizon Oil Spill. *Journal of Environmental Economics and Management* 91: 26–45.

Gellman, J., M. Walls, and M. Wibbenmeyer. 2022. Wildfire, Smoke, and Outdoor Recreation in the Western United States. *Forest Policy and Economics* 134: 102619.

Giglio, L., W. Schroeder, and C.O. Justice. 2016. The Collection 6 MODIS Active Fire Detection Algorithm and Fire Products. *Remote Sensing of the Environment* 178: 31–41.

Goldman, E., A, Wu, K. Kohn, and A, Bartuska. 2022. Visualizing Federal Funding for Wildfire Management and Response. Federation of American Scientists. `https://fas.org/blogs/sciencepolicy/visualizing-federal-funding-for-wildfire-management-and-response`.

Greene, W.H. 2012. *Econometric Analysis* (Seventh edition). Pearson.

Heft-Neal, S., A. Driscoll, W. Yang, G. Shaw, and M. Burke. 2022. Associations Between Wildfire Smoke Exposure During Pregnancy and Risk of Preterm Birth in California. *Environmental Research* 203: 111872.

Heft-Neal, S., C.F. Gould, M. Childs, M.V. Kiang, K. Nadeau, M. Duggan, E. Bendavid, E., and M. Burke. 2023. Behavior Mediates the Health Effects of Extreme Wildfire Smoke Events (No. w30969). National Bureau of Economic Research.

Jones, B. 2017. Are We Underestimating the Economic Costs of Wildfire Smoke? An Investigation Using the Life Satisfaction Approach. *Journal of Forest Economics* 27: 80–90.

Kochi, I., G.H. Donovan, P.A., Champ, and J.B. Loomis. 2010. The Economic Cost of Adverse Health Effects from Wildfire-Smoke Exposure: A Review. *International Journal of Wildland Fire* 19(7): 803–17.

Kolstoe, S., and T.A. Cameron. 2017. The Non-Market Value of Birding Sites and the Marginal Value of Additional Species: Biodiversity in a Random Utility Model of Site Choice by eBird Members. *Ecological Economics*, 137, 1–12.

Korrick, S.A., L.M. Neas, D.W. Dockery, D.R. Gold, G.A. Allen, L.B. Hill, K.D. Kimball, B.A. Rosner, and F.E. Speizer. 1998. Effects of Ozone and Other Pollutants on the Pulmonary Function of Adult Hikers. *Environmental Health Perspectives* 106: 93–9.

Lewis, D.J., S.J. Dundas, D. Kling, D.K. Lew, and S.D. Hacker. 2019. The Non-Market Benefits of Early and Partial Gains in Managing Threatened Salmon. *PLOS ONE* 14(8): e0220260.

Loureiro, M. L., M. Alló, and P. Coello. 2022. Hot in Twitter: Assessing the Emotional Impacts of Wildfires with Sentiment Analysis. *Ecological Economics* 200: 107502.

Lupi, F., D.J. Phaneuf, and R.H. von Haefen. 2020. Best Practices for Implementing Recreation Demand Models. *Review of Environmental Economics and Policy* 14(2): 302–23.

McCoy, S.J., and X. Zhao. 2021. Wildfire and Infant Health: A Geospatial Approach to Estimating the Health Impacts of Wildfire Smoke Exposure. *Applied Economics Letters* 28(1): 32–7.

Miller, N., D. Molitor, and E. Zou. 2021. A Causal Concentration-Response Function for Air Pollution: Evidence from Wildfire Smoke. Working paper, College of Business, University of Illinois, Urbana-Champaign.

Parsons, G.R. 2017. Travel Cost Models. In *A Primer on Nonmarket Valuation*, 2nd ed., edited by P.A. Champ, K.J. Boyle, and T.C. Brown, 269–329. New York: Springer.

Parthum, B., and P. Christensen. 2022. A Market for Snow: Modeling Winter Recreation Patterns Under Current and Future Climate. *Journal of Environmental Economics and Management* 113: 102637.

Preisler, H.K., D. Schweizer, R. Cisneros, T. Procter, M. Ruminski, and L. Tarnay. 2015. A Statistical Model for Determining Impact of Wildland Fires on Particulate Matter (PM2.5) in Central California Aided by Satellite Imagery of Smoke. *Environmental Pollution*: 205: 340–9.

Reid, C.E., M. Brauer, F.H. Johnston, M. Jerrett, J.R. Balmes, and C.T. Elliott. 2016. Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environmental Health Perspectives* 124: 1334–43.

Richardson, L.A., P.A. Champ, and J.B. Loomis. 2012. The Hidden Cost of Wildfires: Economic Valuation of Health Effects of Wildfire Smoke Exposure in Southern California. *Journal of Forest Economics* 18(1): 14–35.

Richardson, L.A., P.A. Champ, and J.B. Loomis. 2013. Valuing Morbidity from Wildfire Smoke Exposure: A Comparison of Revealed and Stated Preference Techniques. *Land Economics* 89(1): 76–100.

Schroeder, W., M. Ruminski, I. Csiszar, L. Giglio, E. Prins, C. Schmidt, and J. Morisette. 2008. Validation Analyses of an Operational Fire Monitoring Product: The Hazard Mapping System. *International Journal of Remote Sensing* 29(20): 6059–66.

Smith, J.W., E.J. Wilkins, and Y.F. Leung. 2019. Attendance Trends Threaten Future Operations of America's State Park Systems. *Proceedings of the National Academy of Sciences* 116(26): 12775–80.

Terza, J.V. 2009. Parametric Nonlinear Regression with Endogenous Switching. *Econometric Reviews* 28(6): 555–80.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Walls, M., C. Wichman, and K. Ankney. 2018. *Nature-Based Recreation: Understanding Campsite Reservations in National Parks*. Washington, DC: Resources for the Future.

Wen, J., and Burke, M. 2022. Lower Test Scores from Wildfire Smoke Exposure. *Nature Sustainability* 5(11): 947–55.

Westerling, A.L. 2016. Increasing Western US Forest Wildfire Activity: Sensitivity to Changes in the Timing of Spring. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1696): 20150178.

Westerling, A.L. 2018. *Wildfire Simulations for California's Fourth Climate Change Assessment: Projecting Changes in Extreme Wildfire Events with a Warming Climate: A Report for California's Fourth Climate Change Assessment* (p. 57). Sacramento, CA: California Energy Commission.

Williams, A.P., J.T. Abatzoglou, A. Gershunov, J. Guzman-Morales, D.A. Bishop, J.K., Balch, and D.P. Lettenmaier. 2019. Observed Impacts of Anthropogenic Climate Change on Wildfire in California. *Earth's Future* 7(8): 892–910.

Wooldridge, J.M. 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources* 50(2): 420–45.

Appendix A: Additional figures

Figure A1: Recreation.gov Web Interface

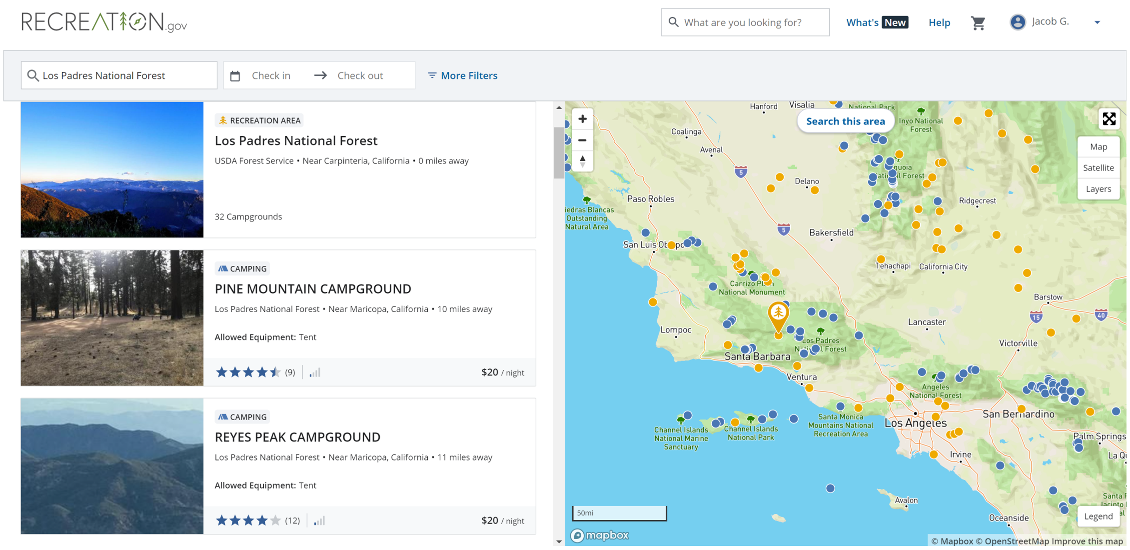


Figure A2: Automobile Route from Santa Barbara, California to Yosemite National Park

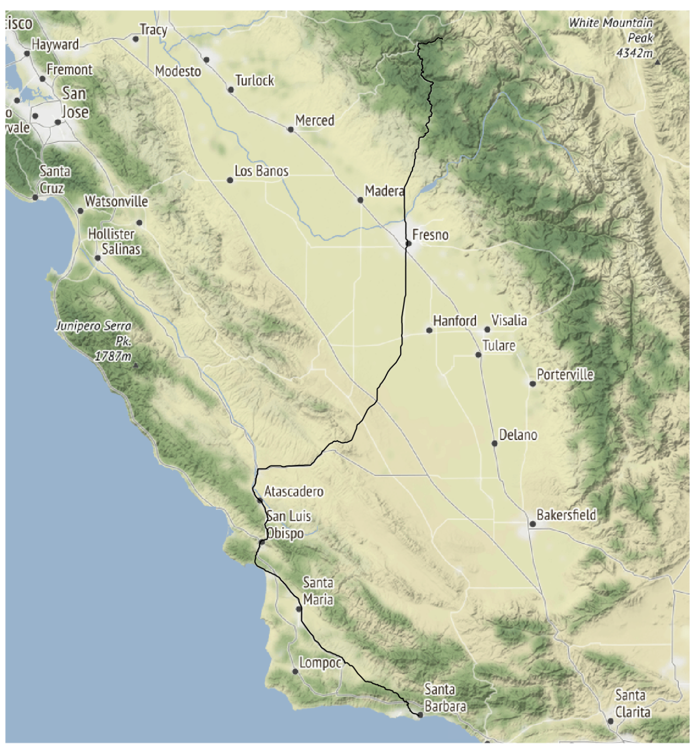
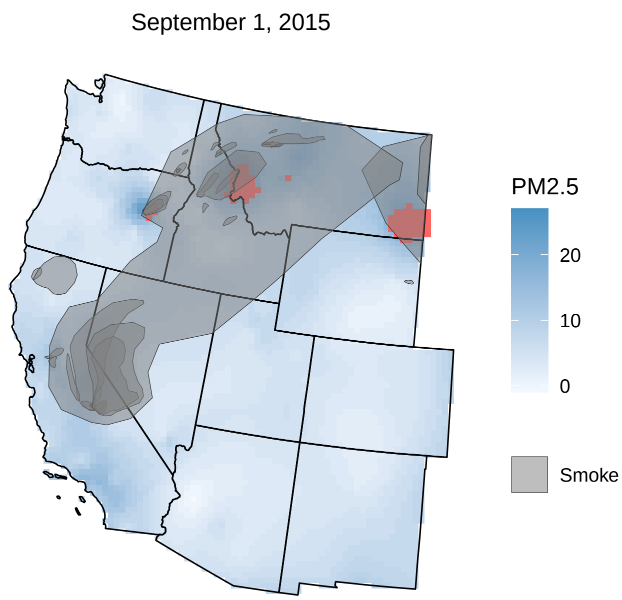


Figure A3: NOAA Smoke Plumes and $PM_{2.5}$.



Note: Red areas are affected by smoke and poor air quality.

Figure A5: Fire Detection Points and Fire Perimeters

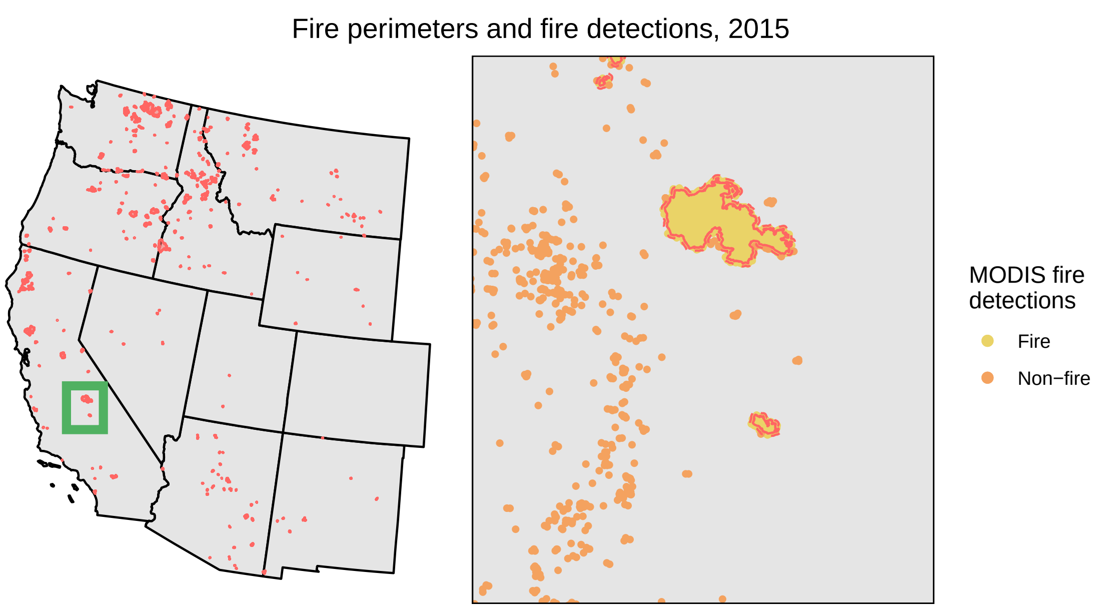


Figure A6: Map of Campgrounds in Dataset

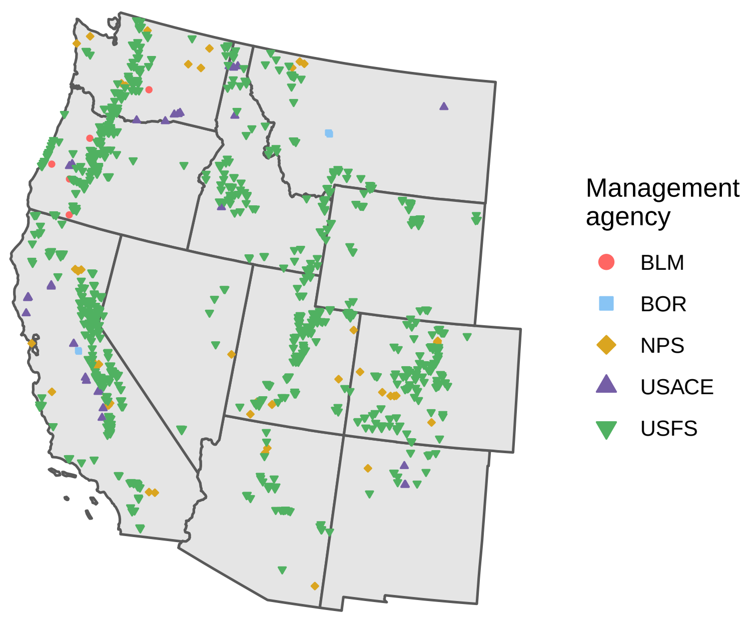


Table A1: Most Visited Federally Managed Campgrounds

Campground	Recreation area	State	Agency	Annual average campers
Upper Pines	Yosemite NP	CA	NPS	99,820
Mather	Grand Canyon NP	AZ	NPS	59,196
Watchman	Zion NP	UT	NPS	49,389
Serrano	Big Bear, San Bernardino NF	CA	USFS	46,610
Pinecrest	Summit RD, Stanislaus NF	CA	USFS	36,576
Fallen Leaf	Lake Tahoe Basin	CA	USFS	32,966
Lodgepole	Sequoia And Kings Canyon NP	CA	NPS	30,634
North Pines	Yosemite NP	CA	NPS	26,883
Moraine Park	Rocky Mountain NP	CO	NPS	25,884
Lower Pines	Yosemite NP	CA	NPS	25,644
Wawona	Yosemite NP	CA	NPS	25,407
Hodgdon Meadow	Yosemite NP	CA	NPS	24,746
Pinnacles	Pinnacles NP	CA	NPS	24,210
Crane Flat	Yosemite NP	CA	NPS	23,844
Indian Cove	Joshua Tree NP	CA	NPS	23,376
Dogwood	Arrow Head, San Bernardino NF	CA	USFS	21,540
Acorn	New Hogan Lake	CA	USACE	21,164
Black Rock	Joshua Tree NP	CA	NPS	19,888
Kalaloch	Olympic NP	WA	NPS	18,105
Dinkey Creek	High Sierra RD, Sierra NF	CA	USFS	16,294
Logger	Truckee RD, Tahoe NF	CA	USFS	16,253
Diamond Lake	Diamond Lake RD, Umpqua NF	OR	USFS	15,683
Kyen	Lake Mendocino	CA	USACE	15,015
Dorst Creek	Sequoia And Kings Canyon NP	CA	NPS	14,435
North Rim	Grand Canyon NP	AZ	NPS	13,898
Ohanapecosh	Mount Rainier NP	WA	NPS	13,889
Devils Garden	Arches NP	UT	NPS	13,138
Oh Ridge	Mono Lake RD, Inyo NF	CA	USFS	13,063
Fish Creek	Glacier NP	MT	NPS	12,434
Manzanita Lake	Lassen Volcanic NP	CA	NPS	12,379

Figure A7: Campground Occupancy Rates Following a Bimodal Distribution Both on the Date of Arrival and One Week in Advance

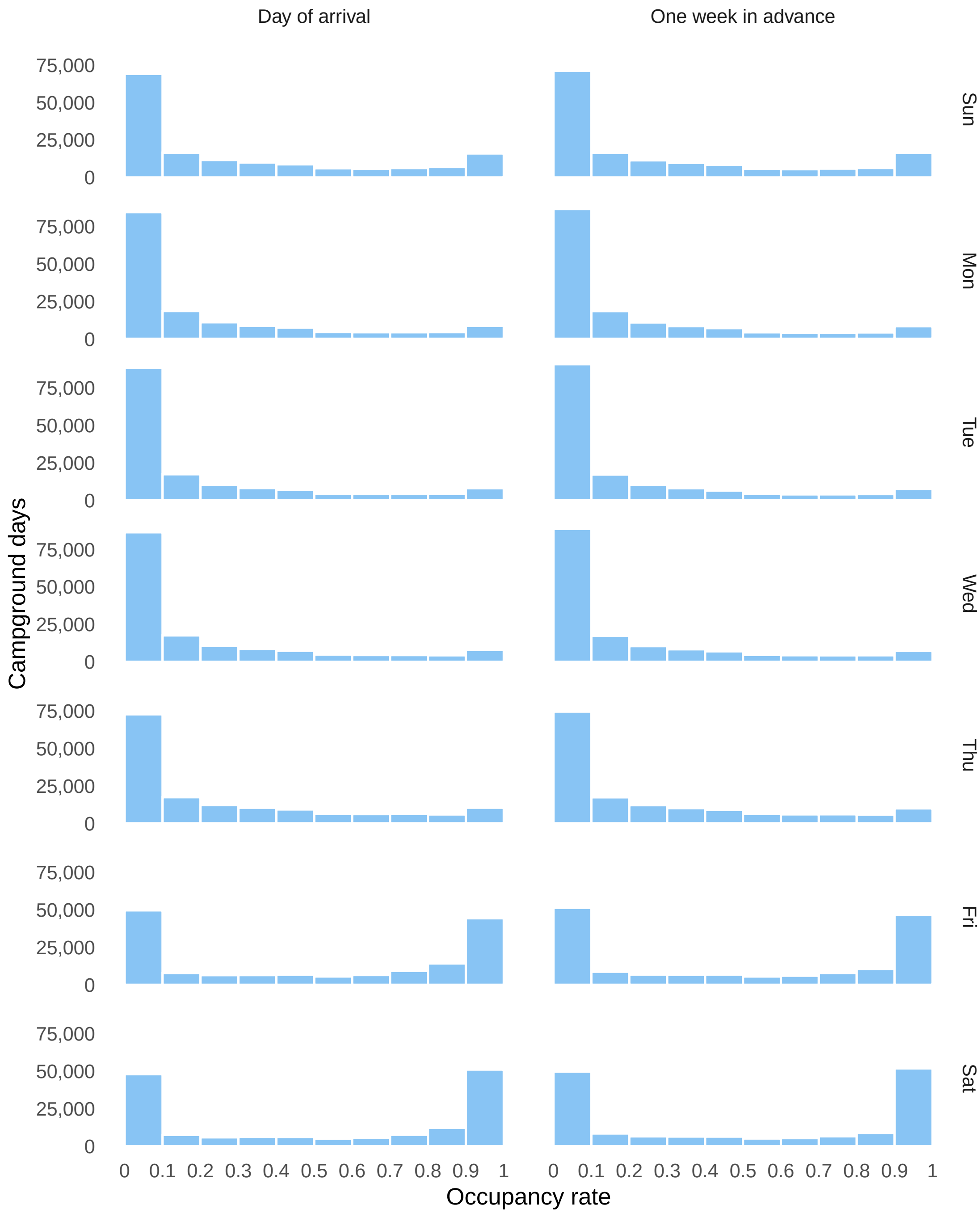


Figure A8: Fitted $\mathbb{P}(R_{ijt} = 1)$ for Reservations Made Earlier Than One Week from Model (4)

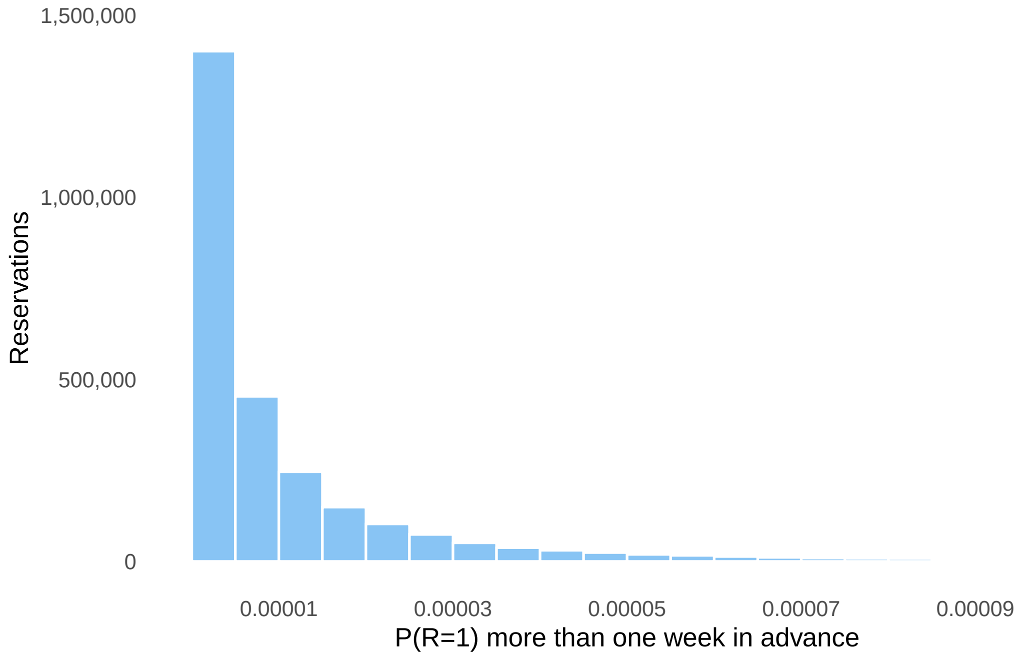


Figure A9: Fitted $\tilde{\varepsilon}_{ijt}$ for Reservations Made Earlier Than One Week from Model (4)

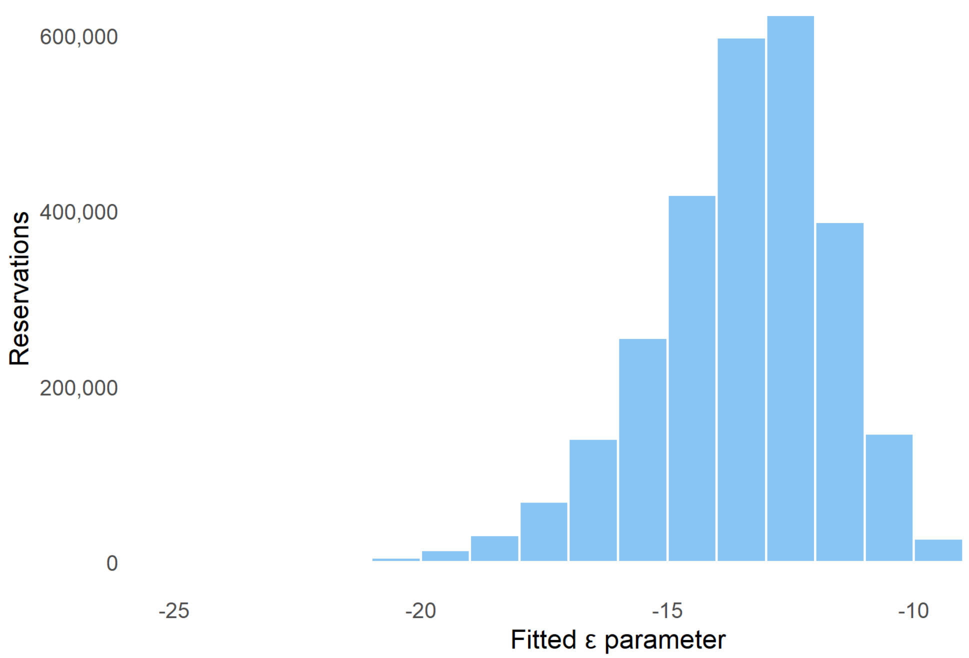


Figure A10: Relationship Between Control Function $\tilde{\varepsilon}_{ijt}$ and Travel Cost Using Model (4) of Table 4 Showing Correlation Between Preferences and Travel Cost in the Selected Sample of Reservers

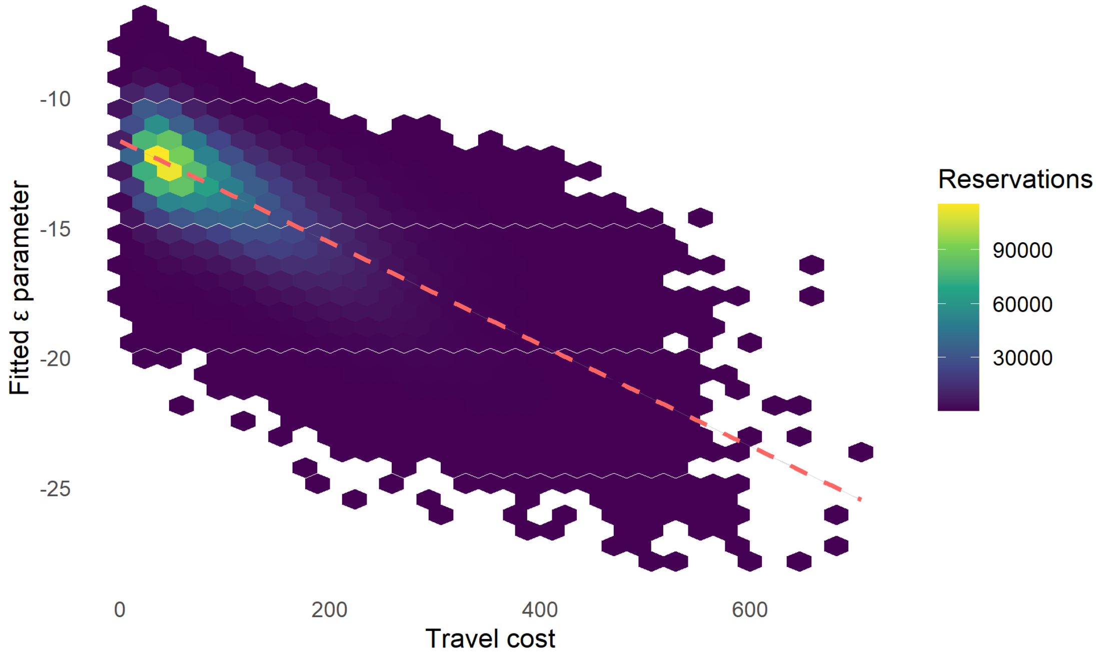


Table A2: Regression of a Binary Variable Equal to 1 if a Campground Is Smoke Affected on a Given Day, as a Function of the Number of Smoke Days in the Week Leading Up to It

	1{Campground is smoke-affected}
Intercept	0.0000
	(0.0001)
1{Smoke days in week of arrival = 1}	0.2013**
	(0.0005)
1{Smoke days in week of arrival = 2}	0.3009**
	(0.0008)
1{Smoke days in week of arrival = 3}	0.3798**
	(0.0010)
1{Smoke days in week of arrival = 4}	0.4850**
	(0.0012)
1{Smoke days in week of arrival = 5}	0.6108**
	(0.0013)
1{Smoke days in week of arrival = 6}	0.7390**
	(0.0016)
1{Smoke days in week of arrival = 7}	1.0000**
	(0.0018)
N	1,528,470

Notes: * $p < 0.05$, ** $p < 0.01$.

Figure A11: Total Estimated Welfare Losses and Proportion of Visits Affected by Smoke per Year

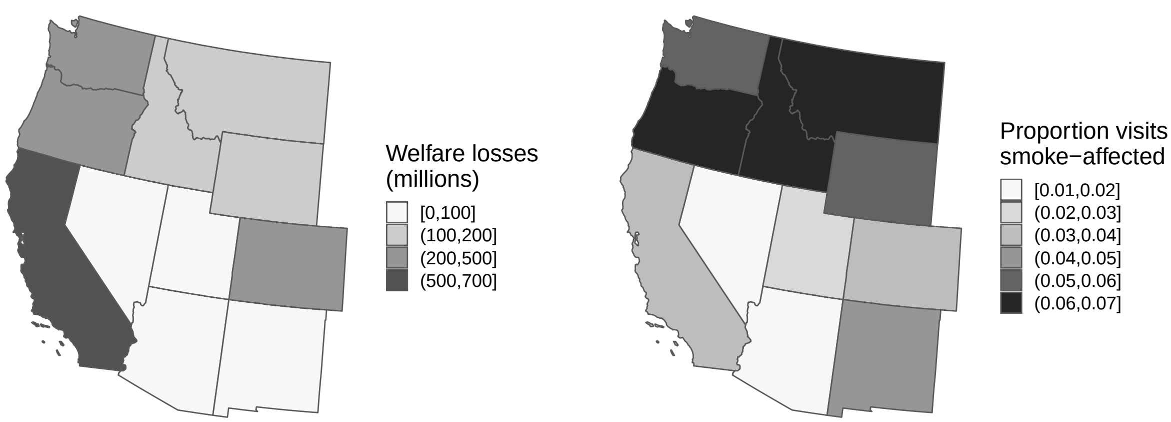


Table A3: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Heterogeneity by Smoke Days in Week Before Arrival

	(1)	(2)	(3)	(4)
Travel cost (dollars)	-0.0025**	-0.0024**	-0.0024**	-0.0024**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-10.5524**	-11.6117**	-11.5449**	-7.4215**
	(0.9020)	(2.3964)	(2.4152)	(0.7843)
High temp. (degrees C)	0.0202**	0.0289**	0.0289**	0.0302**
	(0.0044)	(0.0023)	(0.0023)	(0.0021)
Low temp. (degrees C)	-0.0031	-0.0183**	-0.0189**	-0.0226**
	(0.0057)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0043**	-0.0060**	-0.0061**	-0.0057**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0027	-0.0342**	-0.0352**	-0.0368**
	(0.0255)	(0.0124)	(0.0124)	(0.0126)
Smoke days = 1	0.0158	-0.0718**	-0.0575*	-0.0776**
	(0.0268)	(0.0247)	(0.0246)	(0.0201)
Smoke days = 2	-0.1521**	-0.2164**	-0.1975**	-0.2217**
	(0.0436)	(0.0427)	(0.0416)	(0.0339)
Smoke days = 3	-0.2257**	-0.3050**	-0.2862**	-0.3182**
	(0.0410)	(0.0441)	(0.0437)	(0.0357)
Smoke days = 4	-0.4418**	-0.4792**	-0.4506**	-0.5066**
	(0.0472)	(0.0511)	(0.0502)	(0.0447)
Smoke days = 5	-0.5737**	-0.6032**	-0.5779**	-0.6583**
	(0.0448)	(0.0560)	(0.0551)	(0.0488)
Smoke days = 6	-0.7121**	-0.7612**	-0.7444**	-0.8348**
	(0.0603)	(0.0669)	(0.0669)	(0.0637)
Smoke days = 7	-1.0022**	-1.0065**	-0.9868**	-1.0481**
	(0.0660)	(0.0939)	(0.0922)	(0.0908)
WTP: 1 smoke day	-6.31	30.11*	23.87*	31.96**
	(10.45)	(11.92)	(11.38)	(9.66)
WTP: 2 smoke days	60.79**	90.8**	82.03**	91.32**
	(21.15)	(24.32)	(22.91)	(20.36)
WTP: 3 smoke days	90.26**	127.98**	118.9**	131.09**
	(22.07)	(27.66)	(26.04)	(23.07)
WTP: 4 smoke days	176.63**	201.07**	187.15**	208.68**
	(31.73)	(33.26)	(31.01)	(29.87)
WTP: 5 smoke days	229.38**	253.09**	240.06**	271.19**
	(38.74)	(39.86)	(36.75)	(34.92)
WTP: 6 smoke days	284.7**	319.4**	309.19**	343.87**
	(46.09)	(50.3)	(47.28)	(47.41)
WTP: 7 smoke days	400.7**	422.33**	409.86**	431.74**
	(59.86)	(73.41)	(68.21)	(68.64)
N	2,723,034	2,691,655	2,691,655	2,688,739
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix B: Reservations close to arrival

This paper focuses on the cancellation decisions of visitors who reserve ahead of time, before smoke conditions are known, and decide whether to cancel close to the arrival date, after site conditions are realized. Figure 3 illustrates the timing of decisions in the main analysis. We focus on this structure for several reasons. First, most reservations are made ahead of time. Figure 1 shows that, although a plurality of reservations are made within a week of arrival, the majority are made earlier. Second, by the time smoke conditions are known, many campgrounds are either fully booked or completely empty, which limits the variation needed to identify changes in campground activity due to wildfire smoke. Figure A7 depicts this bimodal distribution. Congested and empty campgrounds both prevent proper measurement of changes in recreation activity due to smoke. When campgrounds are completely booked, logistic regression would underestimate the latent demand for recreation on nonsmoke days because occupancy meets a binding constraint; this analysis would lead to an underestimate of the coefficient on smoke. When campgrounds are empty on nonsmoke days, there is similarly no identifying variation. We focus on cancellations because, once a visitor holds a reservation, they may always cancel it and do not face constraints.[31]

Still, we could have measured decisions for visitors who make new reservations close to the arrival date, when they are likely aware of smoke conditions. In this section, we report results for a zonal travel cost model of new reservations close to the arrival date. We restrict the data to reservations made within a week of arrival during the months of May–September and over the years 2010–2017. We also limit attention to trips coming from within 650 km (400 miles), as described in Section 2.6. Last, we exclude new reservations that were also cancelled in the same week. These restrictions result in 693,501 same-week reservations. We aggregate these reservations for a zonal estimation as described in Section 3.1 but for same-week rather than early reservations.

Figure B1 shows how reservation rates vary by travel cost and wildfire smoke conditions. Reservation rates are much higher at lower levels of travel cost. Before controlling for other

[31]For a discussion of site substitution, refer to Appendix C. Users tend not to cancel and rebook for the same choice occasion. In addition, smoke conditions are spatially and temporally correlated among choice sets, meaning variation is low in differences in smoke-related disutility among choice alternatives.

observable and unobservable factors, Figure B1 shows that raw reservation rates are actually higher on days with smoke, likely because wildfire season overlaps with popular camping times, such as Independence Day and Labor Day. Therefore, fixed effects for location and seasonality are likely to be important.

Figure B1: Reservation Rate Within One Week of Arrival

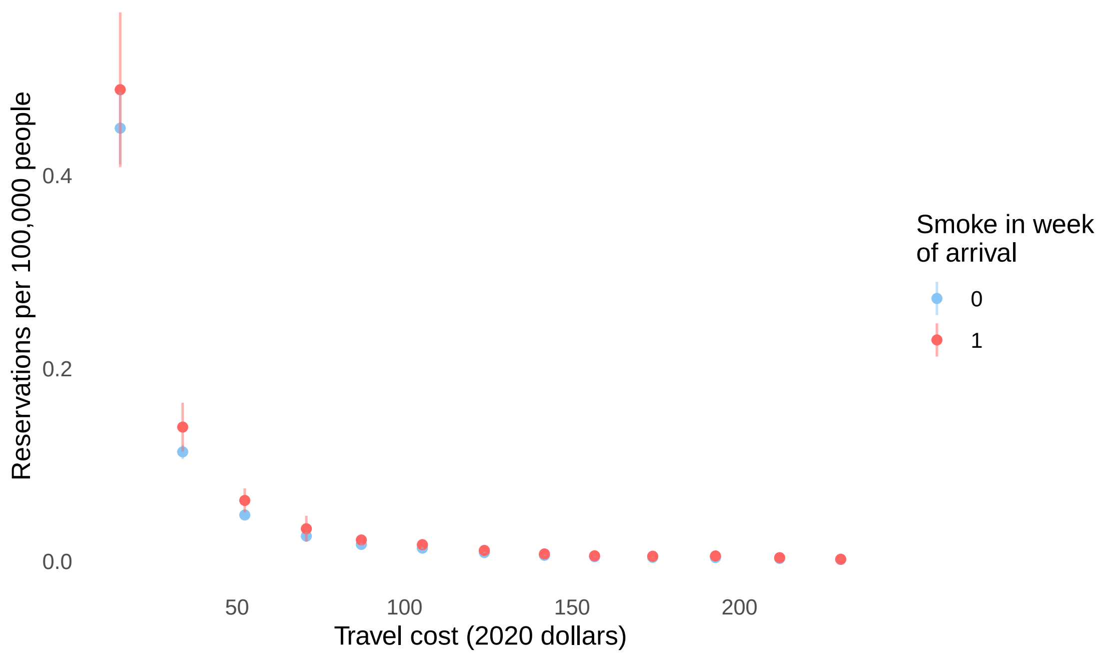


Table B1 reports results for estimation of the reservation likelihood within one week, $\mathbb{P}(R_{ijt} = 1)$, using the zonal maximum likelihood function of Equation 6. In all estimations, we use frequency weights for the observations because a single row of data might represent, for example, 20 reservers or 2.3 million nonreservers. In Column 1, we display results without controlling for campground or seasonal fixed effects. As suggested by Figure B1, users are unconditionally more likely to reserve for smoke-affected dates, yielding an unexpectedly positive coefficient on smoke. Columns 2–4 add fixed effects, which yield the expected sign for the smoke coefficient. The results in Columns 2–4 indicate a WTP to avoid smoke of \$1.45–1.65 per person per trip. Given previous discussion of congested and empty campgrounds, we believe these estimates are less plausible than our main set of results.

Table B1: $\mathbb{P}(R_{ijt} = 1)$ for Reservations Within One Week

	(1)	(2)	(3)	(4)
Smoke in week of arrival	0.1382**	-0.0434**	-0.0406**	-0.0460**
	(0.0021)	(0.0125)	(0.0104)	(0.0091)
Travel cost (dollars)	-0.0241**	-0.0279**	-0.0279**	-0.0279**
	(0.0000)	(0.0017)	(0.0017)	(0.0017)
Inv. distance to wildfire (km^{-1})	-0.6732**	-2.0485**	-2.0798**	-1.9822**
	(0.0310)	(0.3046)	(0.3029)	(0.2809)
High temp. (degrees C)	0.0602**	0.0074**	0.0075**	0.0079**
	(0.0002)	(0.0012)	(0.0011)	(0.0010)
Low temp. (degrees C)	-0.0205**	-0.0044**	-0.0039**	-0.0047**
	(0.0002)	(0.0016)	(0.0015)	(0.0014)
Precip. in week of arrival (mm)	-0.0035**	-0.0028**	-0.0027**	-0.0028**
	(0.0001)	(0.0004)	(0.0004)	(0.0003)
N	13,792,677	10,913,738	10,913,738	10,542,160
WTP	-5.73**	1.55**	1.45**	1.65**
	(0.09)	(0.45)	(0.37)	(0.32)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix C: Site substitution

In the main analysis, both the trip-level cancellation decision and the zonal travel cost model imply a binary choice for the representative visitor. Although we could have used a multinomial logit to model site substitution, we discuss in Section 3.1 the practical limitations of that approach and the advantages afforded by a zonal travel cost model. Moreover, the zonal reservation estimation and binary cancellation decision should properly identify the parameters of interest: the marginal disutility of smoke and of expenditure. In this section, we first show that a binary cancellation decision is a realistic representation of the choice that users face. We also discuss the choice to model early reservations in a zonal setting.

We begin by discussing the binary cancellation decision. Substitution after a cancellation is uncommon. The 2,723,940 trips in the estimating dataset have 268,750 cancellations, implying a 9.87 percent raw cancellation rate. Approximately 10.3 percent of users “rebooked,” meaning they made a new reservation for a date within a year of their original scheduled arrival. However, rebookers rarely substitute for the same choice occasion. Only 11 percent of rebookings substituted to a different campground on the same week of arrival; multiplying by 10.3 percent, this implies that only 1.1 percent of all cancellations did so. Intertemporal substitution is more common: 57 percent of all rebookings were for the same or a different campground at a later week. Multiplying by 10.3 percent, this means that 5.8 percent of all cancellations intertemporally substituted.

Because this analysis is concerned with wildfire smoke, we note the smoke status of rebooked visits: 0.9 and 1 percent were smoke affected and rebooked for a different week or the same week, respectively. Multiplying by 10.3 percent, this means that 0.09 percent and 0.1 percent of all cancellations could have ostensibly substituted due to wildfire smoke. We view these substitutions as uncommon. Therefore, modeling cancellations as a binary decision is a reasonable representation of visitors’ choice.

One additional reason not to model site substitution is that smoke conditions are spatially and temporally correlated, which could wash out differences in smoke-related utility between choice alternatives, variation that is needed to properly identify the smoke parameter. Figures C1, C2, and C3 plot a visualization of this correlation for Colorado, Oregon, and California, sorting campgrounds north to south on the vertical axis and plotting days of the year during the summer months on the horizontal axis. Each tile represents a campground day and is colored according to the smoke conditions on that day. These figures reveal that, when one campground is smoke affected, nearby campgrounds also tend to be. Figure C4 also shows this relationship as a histogram for all campground days in the estimating dataset.

Figure C1: Spatial and Temporal Correlation of Smoke in Colorado

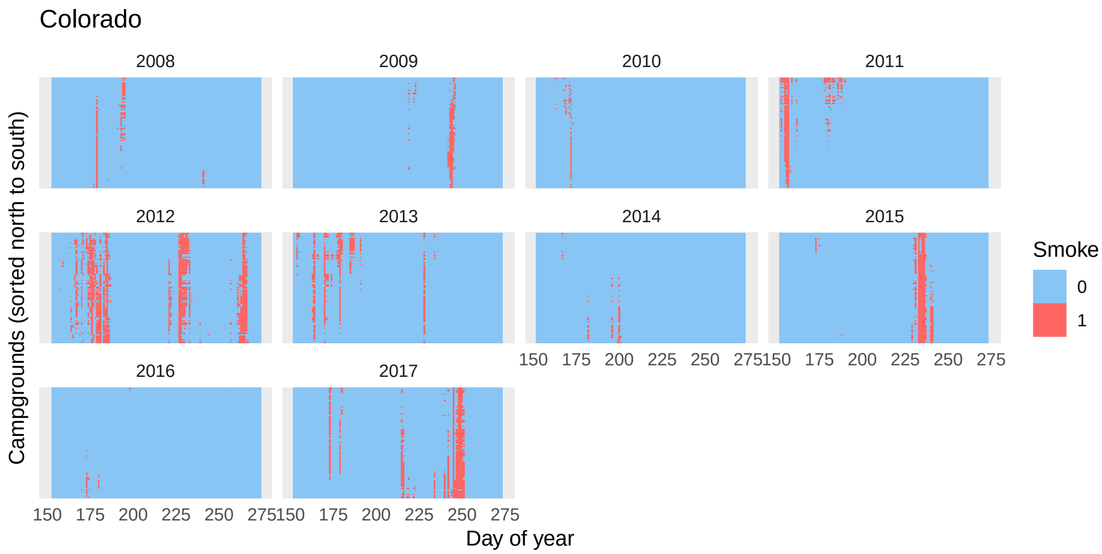


Figure C2: Spatial and Temporal Correlation of Smoke in Oregon

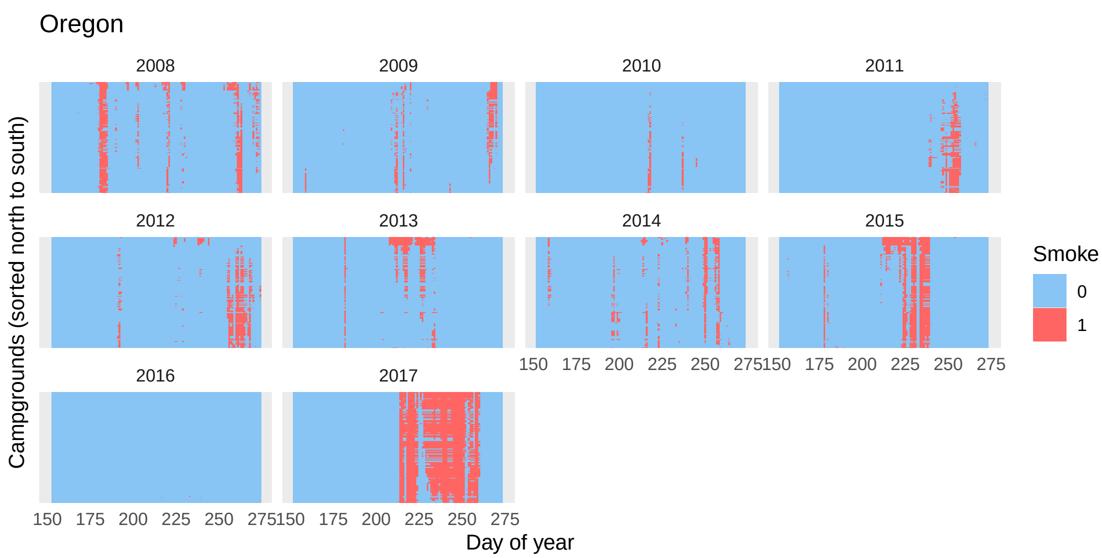


Figure C3: Spatial and Temporal Correlation of Smoke in California

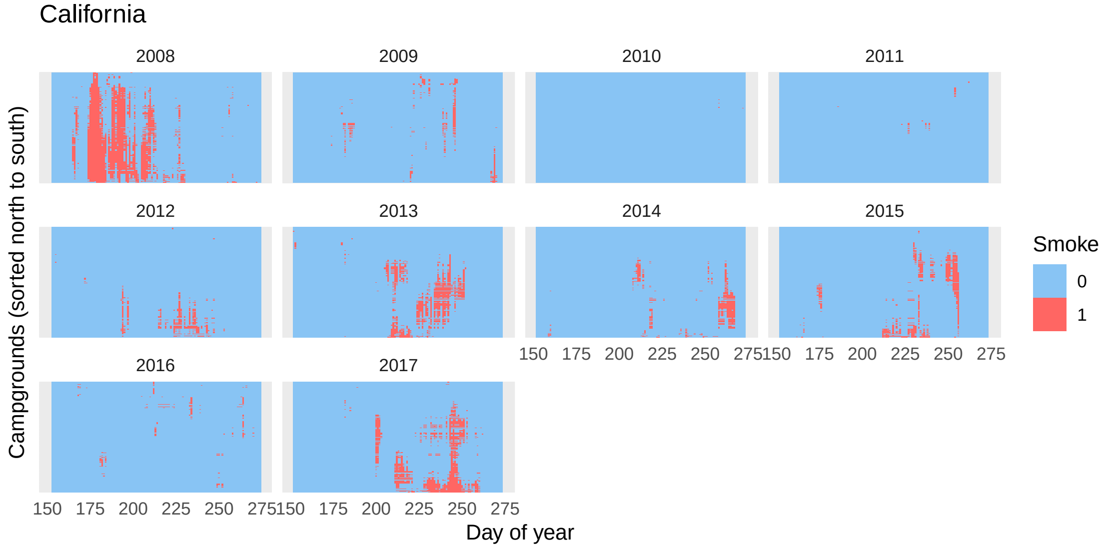
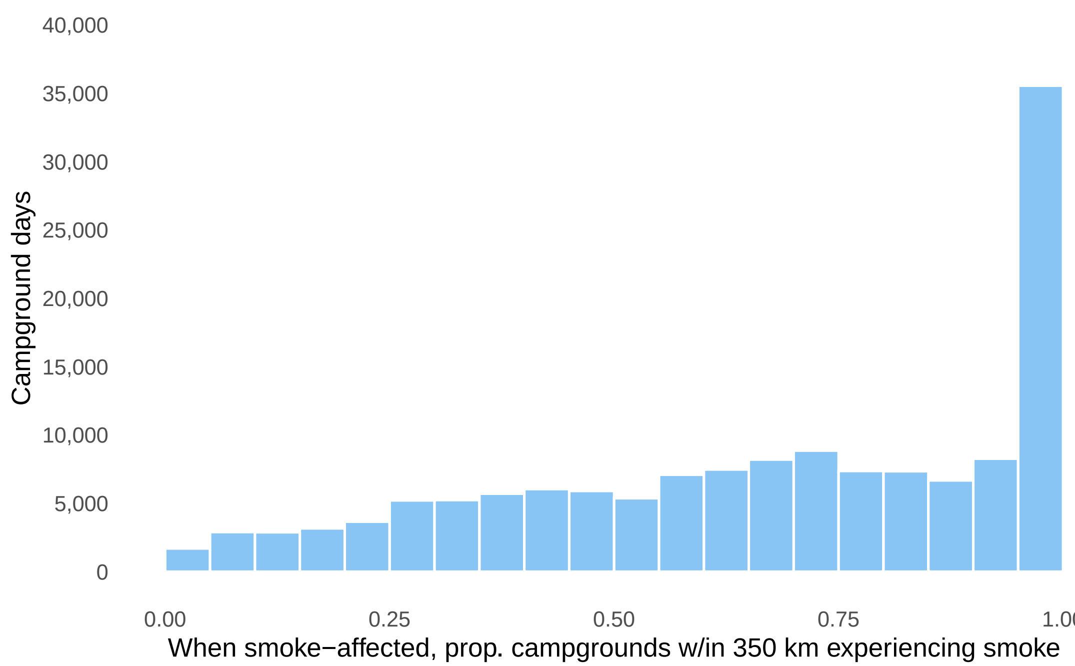


Figure C4: For a Smoke-Affected Campground, Proportion of Campgrounds Within 350 km Experiencing Smoke in the Same Week



The goal of this study is to value the nonmarket damages of wildfire smoke. The parameters of interest to estimate welfare damages are the marginal disutility of wildfire smoke and in expenditure. Both of these arise from the cancellation model, when site conditions become known to individuals. As we discuss in Section 3.1, the main purpose for the reservation estimation is to build the control function that accounts for preferences in the cancellation estimation. These preferences are likely correlated with travel cost in the selected sample, which we explore theoretically in Appendix D and show empirically in Figure A10. We are less concerned with the estimation of smoke in the reservation decision because it occurs before smoke conditions are known. In addition, the binned travel cost zones provide variation to estimate how travel cost affects the likelihood of reservation. Overall, because we are less interested in site substitution for the reservation decision, we argue that the flexible computational advantages afforded by the zonal estimation justify this trade-off. For more discussion, see Section 3.1.

Appendix D: Numerical example of sample selection correction

In Section 3.3, we propose a control function approach to account for unobserved preferences $\tilde{\varepsilon}_{ijt}$ that could bias estimation of $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ if omitted. In this appendix, we provide a numerical example to illustrate the source of this bias, its effect on estimation of willingness to pay (WTP), and correction using a control function. We show that WTP is only biased when preferences for the reservation decisions influence the cancellation decision (i.e., given selection) and the counterfactual cancellation decision of nonreservers is unobserved. Furthermore, the bias operates through correlation between preferences and travel cost: among the selected sample of reservers, those with a high travel cost tend to have had a high taste for the site. This relationship downward biases estimates of the travel cost parameter in the cancellation decision, inflating WTP estimates. Finally, we demonstrate bias correction using the control function for $\tilde{\varepsilon}_{ijt}$ given in Equation 10.

In this numerical example, we simulate the two-stage reservation and cancellation decision using a Monte Carlo of 10,000 random draws. For every iteration, we generate $N = 100,000$ users i, each with a spatial coordinate $(x, y) \in [0, 1] \times [0, 1]$, where x and y are distributed uniformly. In addition, we generate a single site j at a random coordinate $(x, y) \in [0,1] \times [0,1]$, where x and y are again distributed uniformly. User i's travel cost c_{ij} is given by the Euclidean distance from i to j.

Users who reserve far in advance maximize utility based on expected smoke conditions. Define the utility from the reservation as $U_{ij}^R = \alpha_j + \delta c_{ij} + \phi \mathbb{E}[s_j] + \varepsilon_{ij}$, where α_j is an intercept, c_{ij} is the travel cost, s_j is smoke conditions at the site, and ε_{ij} is the individual's preferences from reservation. We will assert arbitrarily that $\alpha_j = 1$, $\delta = -0.8$, and $\phi = -1.6$. Therefore, the true WTP is $\phi/\delta = 2$. Each user's site-specific preference values of ε_{i0} and ε_{ij} are drawn from a type I extreme value distribution. Based on the "time of visitation," expected smoke conditions $\mathbb{E}[s_j]$ are drawn for each user from $\{0.1, 0.2, 0.4\}$ with equal probability. Users will choose to reserve $R_{ij} = 1 \iff U_{ij}^R \geq U_{i0}^R$.

For the cancellation decision, the user decides based on realized smoke conditions. Let the utility from cancellation be $U_{ij}^C = \alpha_j + \delta c_{ij} + \phi s_j + \upsilon_{ij}$. Realized smoke s_j is drawn from $\{0, 1\}$ with $\mathbb{P}(s_j = 1) = 0.25$ for each user to create variation based on the "time of visitation."

We consider two types of error structures υ_{ij} in the cancellation decision. The first is an independent error, $\upsilon_{ij}^{ind} = \frac{1}{\rho}\eta_{ij}$, where $\eta_{ij} \sim$ type I extreme value, which assumes that the user's preferences in the decision are completely unrelated to their choice to have reserved. The second is a dependent error, $\upsilon_{ij}^{dep} = \varepsilon_{ij} + \frac{1}{\rho}\eta_{ij}$, which allows preferences at the time of reservation to affect the decision. We assume $\eta_{ij} \sim$ type I extreme value and arbitrarily set $\rho = 0.7$. Users will cancel $C_{ij} = 1 \iff U_{ij}^C \leq U_{i0}^C$. Because of the differing error structures, we consider two decisions under both υ_{ij}^{ind} and υ_{ij}^{dep}, which we denote as C_{ij}^{ind} and C_{ij}^{dep}, respectively.

The selection issue in the real recreation data arises because we can only observe the cancellation decision for reservers. However, under the Monte Carlo simulation, we can also examine the counterfactual decision of the nonreservers to see if they "would have" cancelled. We show that, even with a dependent error υ_{ij}^{dep}, estimation of $\mathbb{P}(C_{ij} = 0)$ on the full sample (reservers and nonreservers) without observing ε_{ij} will still recover the true WTP because no selection effect exists. That is, the biased estimation of $\mathbb{P}(C_{ij} = 0 | R_{ij} = 1)$ is because ε_{ij} and c_{ij} are correlated in the selected sample, not the full sample.

Table D1: Example of Users' Reservation and Cancellation Decisions

R_{ij}	C_{ij}^{ind}	C_{ij}^{dep}	N
0	0	0	14,060
0	0	1	9,673
0	1	0	17
0	1	1	21,241
1	0	0	29,339
1	0	1	205
1	1	0	9,046
1	1	1	16,419

Table D1 shows an example of users' decisions from one iteration of the Monte Carlo: nonreservers were more likely to cancel with dependent errors, and reservers were less likely to cancel with independent errors. This result is driven by their initial preferences about the site, as reservers have a higher ε_{ij}. Figure D1 illustrates this point by comparing the ε_{ij} of reservers to the total population.

Figure D1: Example Distribution of ε_{ij} for Reservers and All Users

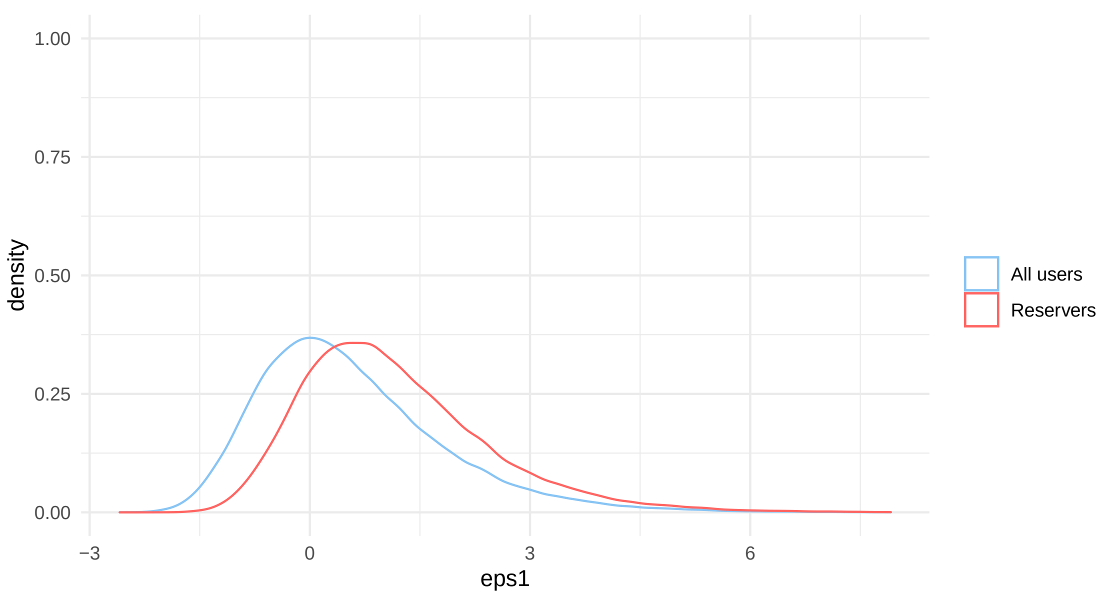


Figure D2 illustrates the effects of selection by contrasting cancellation rates with and without selection effects, illustrating several key points. First, the overall cancellation rate is lower in the presence of selection, as indicated by the intercept of the fitted golden line.

Users who made a reservation had a high initial preference for the site, so they are less likely overall to cancel. Second, the average effect of smoke, which is the distance between the red and blue points, is similar both with and without sample selection effects. Third, the effect of travel cost, which is the slope of the golden fitted line, is attenuated when preferences at the time of reservation affect the cancellation decision. This attenuation illustrates that the selection effect likely operates through positive correlation between ε_{ij} and travel cost.

We can further demonstrate this correlation by regressing distance on ε_{ij} in the full and selected samples. Table D2 shows an example using one draw from the Monte Carlo simulation. Travel cost and distance are correlated for the selected sample but not all users.

Figure D2: Example Cancellation Rate to Illustrate the Effects of Selection

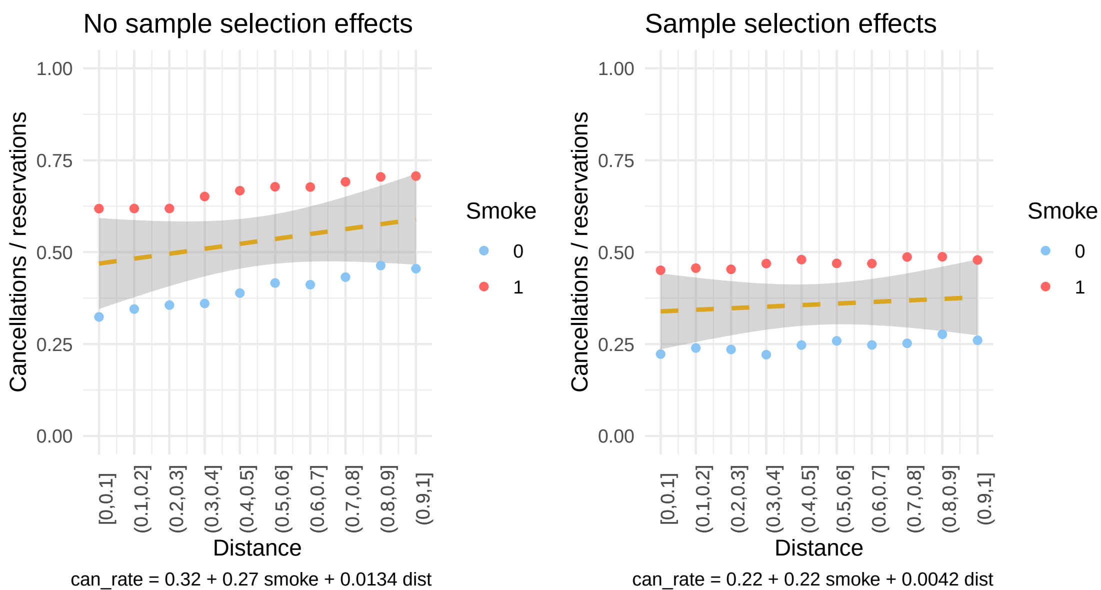


Table D2: Example Regression of Distance on ε_{ij} in the Full and Selected Samples

	(1)	(2)
ε_{ij}	-0.0001	0.0112**
	(0.0006)	(0.0008)
Intercept	0.5236**	0.4904**
	(0.0008)	(0.0014)
Observations	100,000	55,009
R^2	0.0000005	0.004
Users	All users	Reservers
Note:		*p<0.05; **p<0.01.

Next, we show that WTP estimates are only biased under a selected sample and when preferences at the time of reservation affect the cancellation decision. We estimate a logit regression for the reservation and cancellation decisions, varying whether we use the full sample or selected sample of reservers and the dependent error υ_{ij}^{dep} or the independent error υ_{ij}^{ind} for the cancellation decision.

Table D3 shows an example from one iteration of the Monte Carlo simulation. In Column 1, we use the full sample for the reservation decision. In Columns 2 and 3, we estimate the cancellation decision with both errors υ_{ij}^{ind} and υ_{ij}^{dep} but with the full sample. These regressions show that the selection effects would not cause biased estimation if the counterfactual cancellation decision of the nonreservers were known. In Column 4, we estimate the cancellation decision among only the selected sample but with an independent error υ_{ij}^{ind} (i.e., assuming no selection effects). Regression 4 demonstrates that sample selection is not an issue if the user's preferences at the time of reservation are unrelated to their cancellation decision. Finally, Column 5 shows that WTP estimates are biased when the sample is selected and reservation preferences affect the cancellation decision.

Table D3 uses only one draw from the full set of 10,000 random draws. In Table D4, we show the same results over the full set. The logic holds: estimation of WTP is biased only under a selected sample with selection effects.

Table D3: Example Regressions of Reservation and Cancellation Decisions Under Various Samples and Error Structures

	(1)	(2)	(3)	(4)	(5)
Distance	-0.7855**	-0.6303**	-0.4890**	-0.6486**	-0.2411**
	(0.0269)	(0.0274)	(0.0270)	(0.0370)	(0.0398)
$\mathbb{E}$[Smoke]	-1.6406**				
	(0.0514)				
Smoke		-1.1265**	-0.8691**	-1.1116**	-1.0040**
		(0.0155)	(0.0151)	(0.0208)	(0.0206)
Intercept	0.9991**	0.7395**	0.5698**	0.7513**	1.2415**
	(0.0199)	(0.0164)	(0.0161)	(0.0215)	(0.0233)
N	100,000	100,000	100,000	55,009	55,009
Dep. var.	R_{ij}	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Users	All users	All users	All users	Reservers	Reservers
Error	ε_{ij}	υ_{ij}^{ind}	υ_{ij}^{dep}	υ_{ij}^{ind}	υ_{ij}^{dep}
WTP	2.09	1.79	1.78	1.71	4.16

Notes: True WTP = 2. * $p < 0.05$, ** $p < 0.01$.

Table D4: Monte Carlo 10,000 Simulated Regressions of Reservation and Cancellation Decisions

	(1)	(2)	(3)	(4)	(5)
WTP	2.00**	2.01**	2.01**	2.01**	6.70
	(0.10)	(0.11)	(0.14)	(0.15)	(8.52)
Dep. var.	R_{ij}	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Users	All users	All users	All users	Reservers	Reservers
Error	ε_{ij}	υ_{ij}^{ind}	υ_{ij}^{dep}	υ_{ij}^{ind}	υ_{ij}^{dep}

Notes: True WTP = 2. * $p < 0.05$, ** $p < 0.01$.

We next demonstrate the bias correction of the estimand $\tilde{\varepsilon}_{ij}$ derived in Equation 10. We estimate the reservation decision, then use the fitted values of $\bar{V}_{ij}$ to form $\tilde{\varepsilon}_{ij}$. Table D5 shows an example from one draw of the 10,000 simulations. In this example, the smoke coefficient is unaffected by the bias corrector. Instead, the value of the intercept is reduced and the value of the distance coefficient is inflated. In this single random draw, the true

WTP was not exactly recovered. However, over the full set of 10,000 simulations, including $\tilde{\varepsilon}_{ij}$ results in unbiased estimation. Table D6 shows WTP results for the full set. Including the $\tilde{\varepsilon}_{ij}$ estimator resulted in recovery of the true WTP. This result lends support to the use of this bias corrector in the empirical dataset.

Table D5: Example Regression of Cancellation Decision for Reservers Using Bias Correction

	(1)	(2)
Intercept	1.2415**	0.6418**
	(0.0233)	(0.0749)
Smoke	-1.0040**	-1.0057**
	(0.0206)	(0.0206)
Distance	-0.2411**	-0.5480**
	(0.0398)	(0.0540)
$\tilde{\varepsilon}_{ijt}$		-0.6126**
		(0.0729)
N	55,009	55,009
WTP	4.16	1.84
2-step estimator	None	$\tilde{\varepsilon}_{ijt}$

Notes: True WTP = 2. * $p < 0.05$, ** $p < 0.01$.

Table D6: Monte Carlo 10,000 Simulated Regressions Showing Bias Correction

	(1)	(2)
WTP	6.70	2.00**
	(8.52)	(0.25)
2-step estim.	None	$\tilde{\varepsilon}_{ij}$

Notes: True WTP = 2. * $p < 0.05$, ** $p < 0.01$.

This exercise yields several key assumptions. First, the data-generating process asserts that users react the same way to expected and realized smoke. That is, the coefficient for expected and realized smoke is the same. This assumption may not hold for real users; it is reasonable to believe that decision makers may respond differently to expected conditions.

Still, the purpose of $\tilde{\varepsilon}_{ij}$ is to account for selection from the first stage; it should therefore serve as an appropriate control function, regardless of whether the coefficients are identical between stages.

The second key assumption is that the decision maker selects from a single choice alternative. This setup matches our conceptual framework in Section 3, where we assumed a binary site choice. This structure owes to the use of a zonal travel cost model for the first-stage reservation estimation. For an extended treatment of this matter, see Appendix C.

Appendix E: Bootstrapped standard errors for $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$

In Section 4, we used a two-stage sample selection correction to estimate $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$. Wooldridge (2015) recommends that researchers bootstrap standard errors when estimating two-stage control functions. Because we cluster standard errors at the campground level, our bootstrap follows the process outlined by Cameron and Miller (2015) in a methods guide for clustered standard errors: for B bootstraps and G clusters, (1) sample with replacement G times from the original sample of clusters, and (2) compute parameter estimates. The estimating dataset contains $G = 999$ clusters. The resampling occurs over entire clusters; in some bootstraps, some clusters will not be represented, whereas some clusters will have all of their observations appear multiple times in the estimating dataset. Cameron and Miller (2015) note that $B = 400$ should be "more than adequate" in most settings.

In this section, we test that the bootstrapped coefficients follow a normal distribution, assessing whether $B = 400$ is adequate. Table E1 reports W statistics from Shapiro-Wilk tests of normality for the smoke and travel cost coefficients from the main estimation of Table 4. We fail to reject the null hypothesis that the bootstrapped smoke and travel cost coefficients follow a normal distribution. These tests imply that 400 bootstraps are adequate for the analysis. Figures E1 and E2 plot the bootstrapped coefficients visually.

Table E1: W Statistics from Shapiro-Wilk Test of Normality for Bootstrapped Coefficients of $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ with Sample Selection Correction (parentheses indicate p values; the null hypothesis is that the coefficients are normally distributed)

	(1)	(2)	(3)	(4)
Smoke in week of arrival	0.996	0.998	0.998	0.994
	(0.450)	(0.979)	(0.852)	(0.084)
Travel cost (dollars)	0.990	0.996	0.995	0.995
	(0.006)	(0.343)	(0.291)	(0.255)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Figure E1: Distribution of Estimated Smoke Coefficient from Models (1)–(4) in Bootstrapped Estimation of $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ with Sample Selection Correction (red line indicates mean)

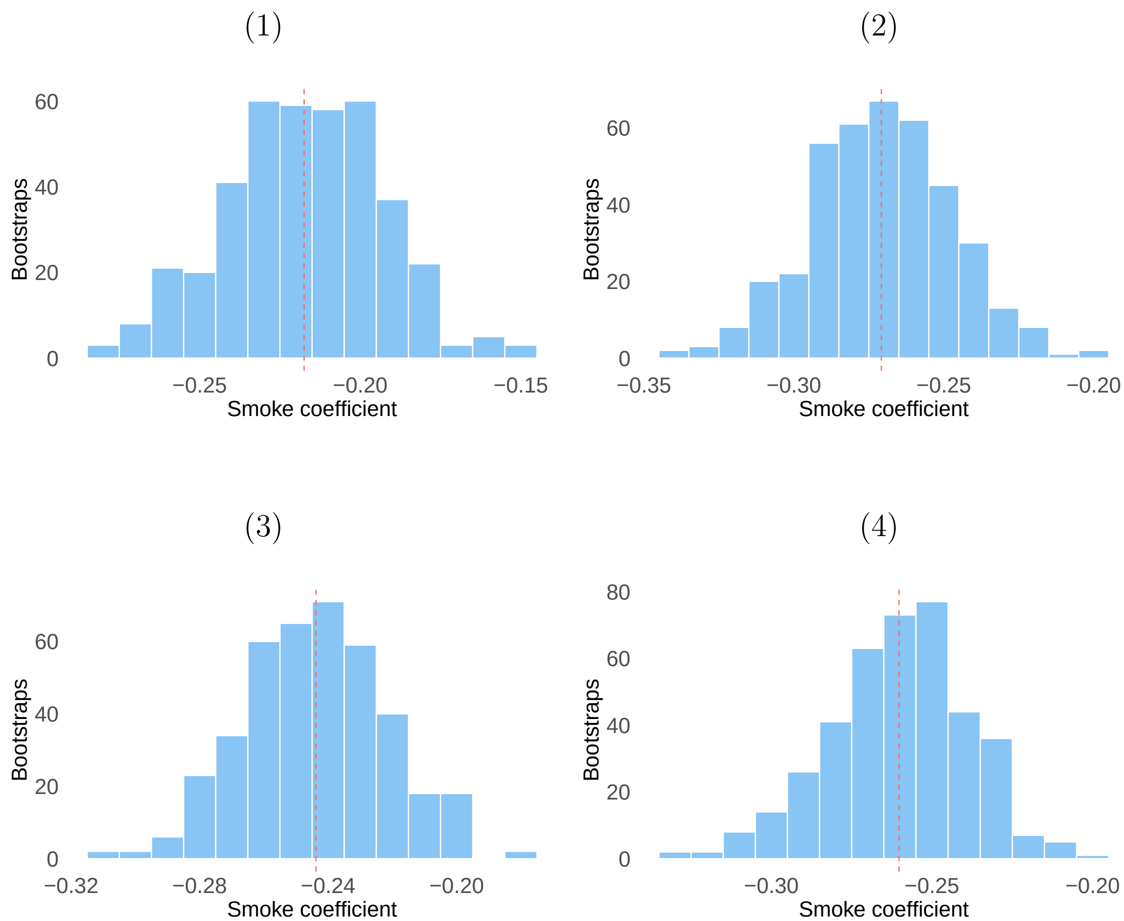
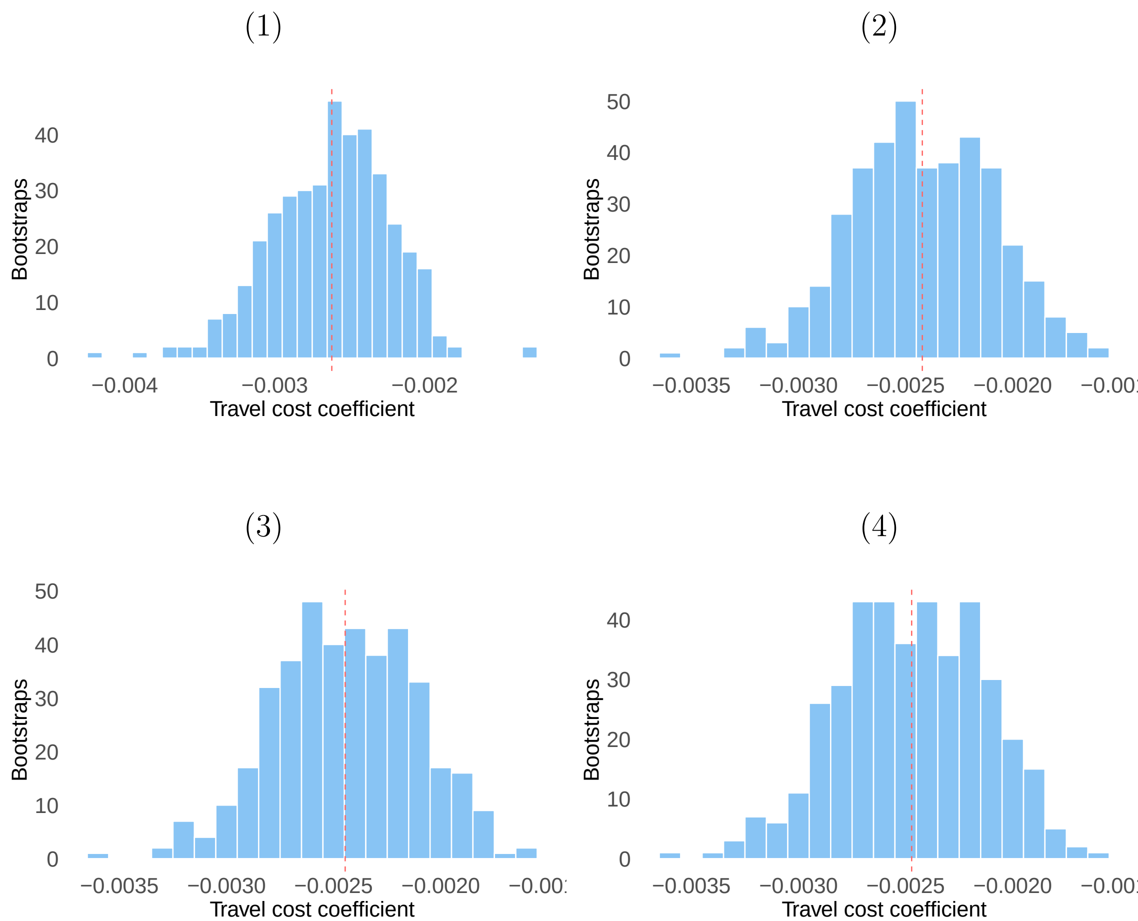


Figure E2: Distribution of Estimated Travel Cost Coefficient from Models (1)–(4) in Bootstrapped Estimation of $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ with Sample Selection Correction (red line indicates mean)



Appendix F: The role of active wildfires

The literature has found that outdoor recreation activity is less responsive to wildfire smoke when an active wildfire is nearby (Cai 2021). In this section, we investigate the role of nearby active wildfires for the parameter estimates. The main estimates of the paper control for inverse distance to active wildfires, but for this analysis, we additionally remove any observation with an active fire close to the campground.

We consider a campground as near to fire on day t if an active wildfire is within 20 km (12 miles), a threshold we have used in previous work (Gellman et al. 2022). Table E2 reports the number of reservations affected by either smoke or fire conditions. When a fire is nearby, smoke- and non-smoke-affected reservations are nearly equal. However, due to the large distances that smoke travels, most smoke-affected reservations are not for campgrounds near an actively burning wildfire.

Table E2: Reservations with Smoke or Fire Conditions in the Estimating Dataset

Smoke in week of arrival	Fire within 20 km	Number of reservations	Percent of sample (%)
0	0	2,356,407	86.5
1	0	322,114	11.8
0	1	24,199	0.9
1	1	21,220	0.8

Table E3 reports results for the cancellation estimation $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ when removing observations with nearby active wildfire. We find welfare damages of \$85 per person per trip due to smoke. By comparison, in the main specification, we estimated \$107 per person per trip. The omission of fire days thus reduced estimated welfare damages by approximately 20 percent. These results are broadly consistent with Cai (2021), which found that outdoor recreation was less responsive to smoke coming from distant wildfires.

Table E3: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Removing Days with Wildfire Within 20 km

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.1678**	-0.2323**	-0.2119**	-0.2116**
	(0.0224)	(0.0195)	(0.0173)	(0.0177)
Travel cost (dollars)	-0.0028**	-0.0024**	-0.0025**	-0.0025**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-9.4566**	-7.7163**	-7.2826**	-6.5691**
	(1.1591)	(1.0196)	(1.0794)	(0.8965)
High temp. (degrees C)	0.0195**	0.0275**	0.0277**	0.0298**
	(0.0043)	(0.0021)	(0.0021)	(0.0021)
Low temp. (degrees C)	-0.0029	-0.0202**	-0.0208**	-0.0249**
	(0.0056)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0043**	-0.0060**	-0.0062**	-0.0059**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0129	-0.0348**	-0.0357**	-0.0377**
	(0.0261)	(0.0121)	(0.0121)	(0.0123)
N	2,677,628	2,645,592	2,645,592	2,642,695
WTP	60.97**	95.08**	85.96**	84.66**
	(12.4)	(13.13)	(12.03)	(12)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix G: Testing the influence of no-shows in cancellations

One may be concerned that some recreationists do not formally cancel their reservation when they decide not to complete a trip. Unreported no-shows threaten the identification of any willingness to pay (WTP) that is based on cancellations, as it could underestimate them. Although most of the campgrounds in the Recreation.gov dataset do not report check-ins or no-shows, a subset do.

In this section, we compare estimates at these select campgrounds with and without including no-shows. We demonstrate that although no-shows are not infrequent at these campgrounds, omitting them does not influence measures of responses to smoke and travel cost. This analysis should mitigate some concern that we understate avoidance behavior.

Just 36 out of 999 campgrounds (3.6 percent) report no-shows, but they represent a large proportion of the reservations used in the cancellation estimation. Of the reservations made more than a week ahead of time, 2,188,444 were at non-no-show facilities (80.3 percent), and 535,496 were at facilities that report no-shows (19.7 percent).

To gauge the importance of no-shows in the cancellation estimation, Figure G1 and Table G1 report the share of all cancellations that are no-shows at each campground. Although most of the overall dataset comprises US Forest Service campgrounds, Table G1 shows that many campgrounds reporting no-shows are managed by the National Park Service and US Army Corps of Engineers. For most of these, no-shows represent less than 15 percent of all cancellations, although they are nearly a third at some campgrounds.

Figure G1: No-Shows as a Proportion of All Cancellations Among Campgrounds Reporting Them

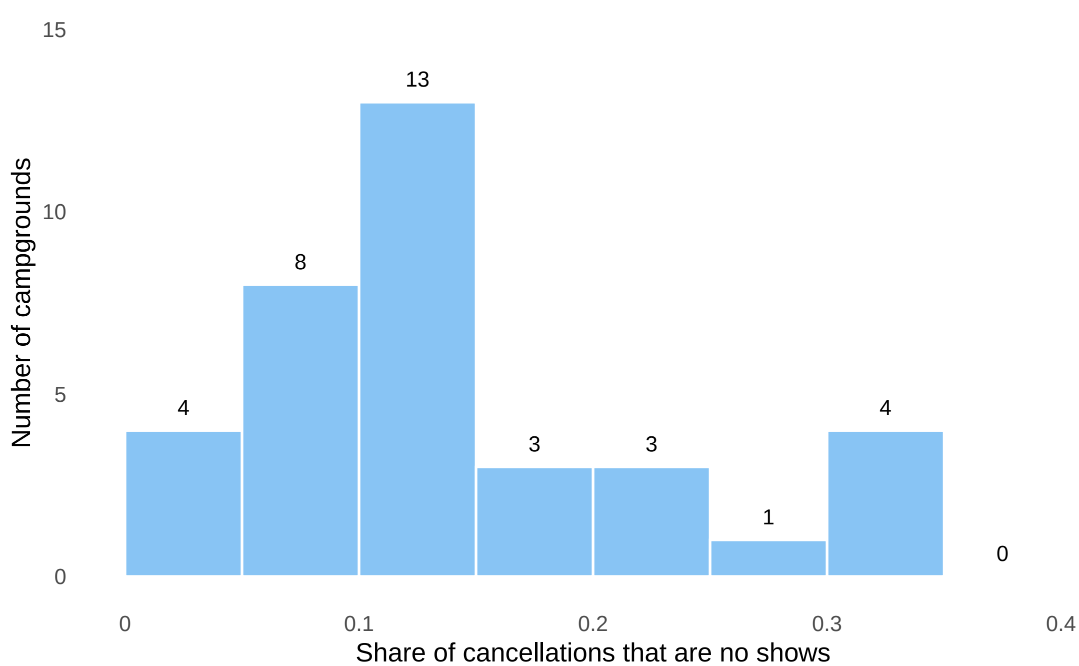


Table G1: Campgrounds Reporting No-Shows

Campground	Recreation area	State	Agency	No-show % of cancellations
Aspenglen Campground	Rocky Mountain National Park	CO	NPS	33.69
Glacier Basin Campground	Rocky Mountain National Park	CO	NPS	32.8
Moraine Park Campground	Rocky Mountain National Park	CO	NPS	32.51
Watchman Campground	Zion National Park	UT	NPS	32.36
Mather Campground	Grand Canyon National Park	AZ	NPS	27.59
Schwarz Park	Dorena Lake	OR	USACE	24.62
Buckhorn	Black Butte Lake	CA	USACE	23.37
Springy Point	Albeni Falls Dam	ID	USACE	20.44
Hood Park	McNary Lock And Dam	WA	USACE	18.51
Hodgdon Meadow	Yosemite National Park	CA	NPS	18.49
Fishhook Park	Ice Harbor Lock	WA	USACE	15.09
Dinkey Creek	High Sierra RD	CA	USFS	14.86
Meeks Bay	Lake Tahoe Basin	CA	USFS	14.03
Serrano	Big Bear	CA	USFS	14.02
Crane Flat	Yosemite National Park	CA	NPS	13.89
Riley Creek	Albeni Falls Dam	ID	USACE	13.83
Dogwood	Arrow Head	CA	USFS	13.24
Wawona	Yosemite National Park	CA	NPS	13.2
Charbonneau Pk	Ice Harbor Lock	WA	USACE	12.74
Pine Meadows Campground	Cottage Grove Lake	OR	USACE	12.65
North Rim Campground	Grand Canyon National Park	AZ	NPS	12.47
Kyen Campground	Lake Mendocino	CA	USACE	10.83
Nevada Beach Campground	Lake Tahoe Basin	CA	USFS	10.81
William Kent Campground	Lake Tahoe Basin	CA	USFS	10.1
Lepage Park	John Day Lock	OR	USACE	9.74
Fish Creek Campground	Glacier National Park	MT	NPS	8.5
Rancheria	High Sierra RD	CA	USFS	7.61
Oh Ridge	Mono Lake RD	CA	USFS	7.19
Deer Creek	High Sierra RD	CA	USFS	7.1
Diamond Lake	Diamond Lake RD	OR	USFS	6.01
Downstream	Fort Peck Project	MT	USACE	5.68
Pinecrest	Summit RD	CA	USFS	5.47
Fallen Leaf Campground	Lake Tahoe Basin	CA	USFS	4.56
Acorn Campground	New Hogan Lake	CA	USACE	4.46
Lodgepole Campground	Sequoia And Kings Canyon National Park	CA	NPS	2.5
Dorst Creek Campground	Sequoia And Kings Canyon National Park	CA	NPS	2.33

We use this subset to demonstrate that unreported no-shows likely do not matter in the full sample. Table G2 tests whether these campgrounds are systematically different than the full sample: they tend to have higher cancellation rates under both smoke and nonsmoke conditions.

However, the estimates in Table G3 should alleviate concerns that no-shows are influential in cancellation estimates. We test four models with the same sets of fixed effects but vary the estimating sample. In Column 1, we include all campgrounds and estimate $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$. Column 2 removes no-shows from the dataset, finding that WTP is unchanged. In Column 3, we allow smoke and travel cost to respond differentially for no-show and non-no-show campgrounds but include no-shows in the dataset. This model shows that no-show and non-no-show campgrounds have different overall measures of WTP. Finally, Column 4 removes no-shows from the dataset. Comparing the WTP of no-show campgrounds with and without the inclusion of no-shows, WTP is virtually unchanged. This analysis should alleviate concerns that no-shows influence the estimate of WTP in the full sample.

Table G2: Cancellation Rate Mean by Campground Type and Smoke

	All campgrounds	Non-no-show campgrounds	No-show campgrounds	t-statistic
Baseline	0.09	0.09	0.13	(9.44)
No. of res.	2,380,606	1,898,955	481,651	
Smoke	0.13	0.12	0.19	(2.02)
No. of res.	343,334	289,489	53,845	
t-statistic	(13.41)	(13.19)	(3.06)	

Notes: The righthand column gives the t-statistic for the difference in mean cancellation rates by campground type in either smoke or nonsmoke conditions. The bottom row gives the t-statistic for the difference in mean cancellation rate for smoke and nonsmoke days among the different campground types.

Table G3: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Testing Effect of No-Shows on Cancellation

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2613**	-0.2659**		
	(0.0215)	(0.0215)		
Smoke x 1(Non-no-show campground)			-0.2608**	-0.2605**
			(0.0188)	(0.0188)
Smoke x 1(No-show campground)			-0.2628**	-0.2846**
			(0.0689)	(0.0736)
Travel cost (dollars)	-0.0024**	-0.0025**		
	(0.0003)	(0.0004)		
Travel cost x 1(Non-no-show campground)			-0.0025**	-0.0025**
			(0.0004)	(0.0004)
Travel cost x 1(No-show campground)			-0.0023**	-0.0024**
			(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-7.8194**	-7.9970**	-7.8180**	-7.9907**
	(0.8239)	(0.8659)	(0.8223)	(0.8631)
High temp. (degrees C)	0.0306**	0.0308**	0.0306**	0.0308**
	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Low temp. (degrees C)	-0.0252**	-0.0254**	-0.0252**	-0.0254**
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0057**	-0.0058**	-0.0057**	-0.0058**
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0370**	-0.0370**	-0.0376**	-0.0373**
	(0.0126)	(0.0131)	(0.0134)	(0.0137)
WTP	107.95**	107.92**		
	(17.14)	(17.48)		
WTP, non-no-show campgrounds			103.82**	103.9**
			(17.39)	(17.75)
WTP, no-show campgrounds			116.66**	119.75**
			(32.1)	(33.5)
No-shows included?	Yes	No	Yes	No
N	2,688,739	2,677,763	2,688,739	2,677,763
Day-of-week FE	Yes	Yes	Yes	Yes
Campground x week-of-year FE	Yes	Yes	Yes	Yes
Campground x year FE	Yes	Yes	Yes	Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix H: Alternative distance thresholds for sample restriction

The main estimates of this paper restrict the estimating sample to reservations from origins within driving distance of a site, which we define as 650 km of one-way driving distance, or approximately 400 miles. Figure 1 shows that this threshold admits approximately 85 percent of the total reservations into the estimation. In this section, we show results from the main estimation using alternative distance thresholds of 350 km (approximately 217 miles) and 950 km (approximately 590 miles).

Figure H1 illustrates how willingness to pay (WTP) estimates increase as the distance threshold is relaxed. Using a restrictive threshold of 350 km, WTP is estimated to be $79 per person per trip; with a wider threshold of 950 km, WTP is estimated to be $140 per person per trip. Tables H1 and H2 show full results for these estimations, which should be compared to the main estimates in Table 4.

WTP is calculated as the ratio of marginal disutility in smoke to that in expenditure (the smoke coefficient divided by the travel cost coefficient). The choice of distance threshold does not alter the estimated smoke coefficient. Instead, the increasing WTP estimates are driven by a decline in the magnitude of the travel cost coefficient as the distance threshold is relaxed. The travel cost coefficient is estimated at -0.0034, -0.0025, and -0.0019 for thresholds of 350 km, 650 km, and 950 km, respectively. In other words, increasing the pool of potential reservers decreases the estimated response to travel cost. This phenomenon could result from including visitors at greater distances who chose not to cancel their reservations.

An additional difference across estimations is the magnitude of the coefficient for the $\tilde{\varepsilon}_{ijt}$ preference parameter, which is estimated at -0.0240, -0.0385, and -0.0390 for 350 km, 650 km, and 950 km, respectively. The magnitude is likely smaller at lower thresholds due to the correlation of preferences with travel cost; removing reservations from larger distances eliminates some visitors with both high travel costs and high preferences. Figure H2 shows that the fitted parameter $\tilde{\varepsilon}_{ijt}$ correlates with travel cost in both the 350 km and 950 km samples.

Figure H1: Summary of Willingness to Pay (WTP) Measures Using Alternative Distance Thresholds for Sample Restriction of 350 km (217 miles) and 950 km (590 miles)

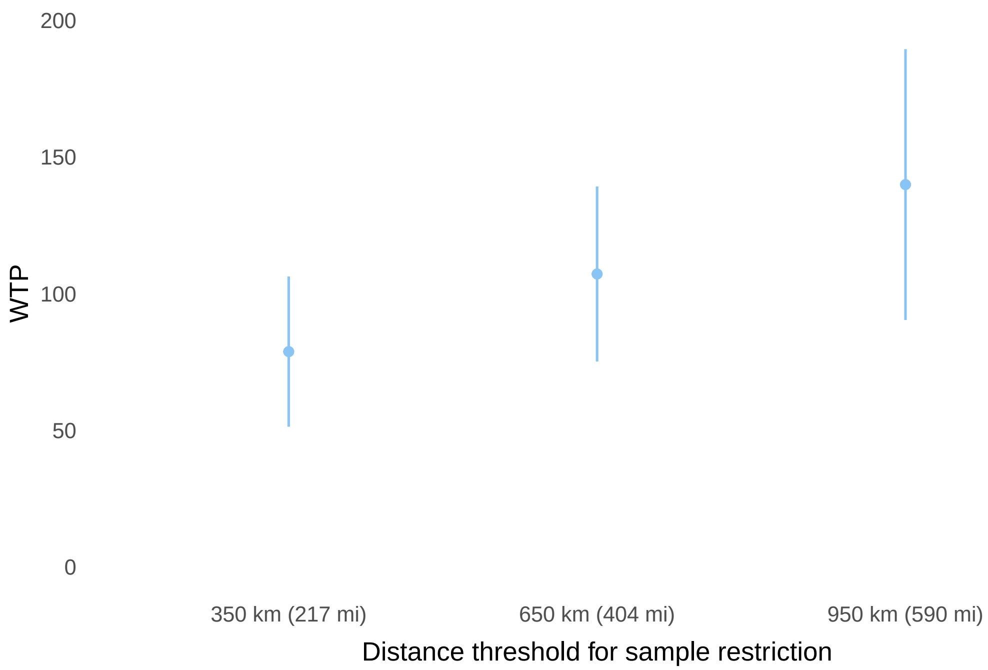


Figure H2: Relationship Between Control Function $\tilde{\varepsilon}_{ijt}$ and Travel Cost Using Model (4), Using a Distance Threshold of (A) 350 km (217 miles) and (B) 950 km (590 miles) (compare to Figure A10)

(A) 350 km. (B) 950 km.

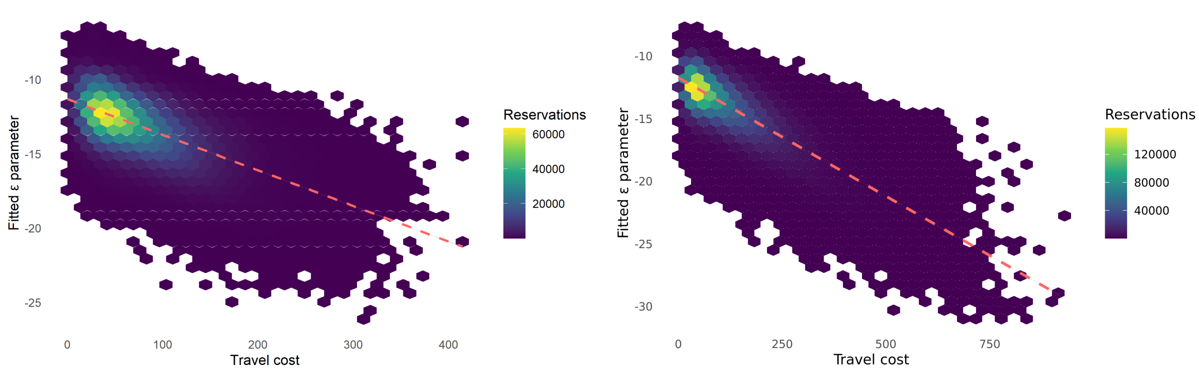


Table H1: $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ Within One Week, Restricting Sample Distance to Within 350 km (217 miles) of Site

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2227**	-0.2641**	-0.2329**	-0.2651**
	(0.0262)	(0.0342)	(0.0324)	(0.0241)
Travel cost (dollars)	-0.0039**	-0.0033**	-0.0033**	-0.0034**
	(0.0006)	(0.0005)	(0.0005)	(0.0005)
Inv. distance to wildfire (km^{-1})	-12.3006**	-14.1542**	-14.0303**	-8.9377**
	(1.0509)	(3.2997)	(3.3356)	(0.8670)
High temp. (degrees C)	0.0210**	0.0323**	0.0325**	0.0331**
	(0.0043)	(0.0027)	(0.0027)	(0.0025)
Low temp. (degrees C)	-0.0007	-0.0209**	-0.0217**	-0.0248**
	(0.0054)	(0.0029)	(0.0029)	(0.0028)
Precip. in week of arrival (mm)	-0.0045**	-0.0065**	-0.0066**	-0.0061**
	(0.0011)	(0.0010)	(0.0010)	(0.0010)
$\tilde{\varepsilon}_{ijt}$	-0.0139	-0.0222	-0.0230	-0.0240
	(0.0344)	(0.0160)	(0.0162)	(0.0165)
N	2,085,985	2,047,894	2,047,894	2,044,062
WTP	57.42**	80.09**	69.86**	78.74**
	(11.65)	(17.61)	(15.7)	(14.02)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Table H2: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within One Week, Restricting Sample Distance to Within 950 km (590 miles) of Site

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2128**	-0.2622**	-0.2354**	-0.2589**
	(0.0238)	(0.0269)	(0.0259)	(0.0206)
Travel cost (dollars)	-0.0021**	-0.0019**	-0.0019**	-0.0019**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-10.9723**	-11.6811**	-11.5690**	-7.5899**
	(0.9010)	(2.2906)	(2.3049)	(0.8101)
High temp. (degrees C)	0.0198**	0.0280**	0.0286**	0.0300**
	(0.0046)	(0.0023)	(0.0022)	(0.0021)
Low temp. (degrees C)	-0.0053	-0.0199**	-0.0210**	-0.0249**
	(0.0058)	(0.0024)	(0.0024)	(0.0024)
Precip. in week of arrival (mm)	-0.0039**	-0.0058**	-0.0060**	-0.0056**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0121	-0.0414**	-0.0424**	-0.0390**
	(0.0243)	(0.0122)	(0.0121)	(0.0128)
N	2,884,364	2,854,171	2,854,171	2,851,414
WTP	100.43**	137.19**	121.83**	139.83**
	(20.95)	(27.77)	(24.69)	(25.27)
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix I: Heterogeneous results within τ days of arrival

The main estimation considers the cancellation decisions of users within $\tau = 7$ days of arrival, where τ is defined in Figure 3. In this section, we explore heterogeneous results for alternative temporal thresholds. We reconstruct the dataset to estimate visitors' probability of cancellation within $\tau = 3$, 5, 7, and 9 days of arrival. The variable of interest is an indicator equal to 1 if a smoke-affected day occurred within the τ day threshold. We consider only standing, uncancelled reservations as of τ days before arrival.

Table I1 reports these results.[32] For $\tau = 3$, 5, 7, and 9 days, we find welfare damages of $137, $129, $107, and $92 per person per trip, respectively. These results are consistent with an information mechanism, which was explored in Section 4.2. For smaller values of τ, the occurrence of one smoke day corresponds to a greater likelihood of smoke on the actual day of arrival. Visitors may have a greater propensity to cancel when observing smoke closer to the date of arrival. The travel cost coefficient is largely stable as τ decreases; greater willingness to pay (WTP) is driven by a growth in the magnitude of the smoke coefficient.

[32] In addition, refer to Table 4 in the main text for results when $\tau = 7$.

Table I1: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within τ Days of Arrival

τ	3 days	5 days	7 days	9 days
Smoke	-0.2961**	-0.2929**	-0.2603**	-0.2281**
	(0.0536)	(0.0241)	(0.0218)	(0.0207)
Travel cost (dollars)	-0.0022**	-0.0023**	-0.0025**	-0.0025**
	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Inv. distance to wildfire ($\mathrm{km}^{-}1$)	-12.2894*	-7.9469**	-7.8141**	-7.7057**
	(4.9610)	(0.8285)	(0.7920)	(0.8411)
High temp. (degrees C)	0.0384**	0.0343**	0.0306**	0.0284**
	(0.0026)	(0.0023)	(0.0023)	(0.0020)
Low temp. (degrees C)	-0.0310**	-0.0279**	-0.0252**	-0.0235**
	(0.0029)	(0.0026)	(0.0025)	(0.0023)
Precip. (mm)	-0.0052**	-0.0054**	-0.0057**	-0.0055**
	(0.0009)	(0.0008)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0291*	-0.0353**	-0.0385**	-0.0380**
	(0.0147)	(0.0136)	(0.0106)	(0.0126)
N	2,917,431	2,783,520	2,688,739	2,602,897
WTP	136.73**	128.9**	107.14**	91.69**
	(37.43)	(22.22)	(16.33)	(15.03)
Day-of-week FE	Yes	Yes	Yes	Yes
Campground x week-of-year FE	Yes	Yes	Yes	Yes
Campground x year FE	Yes	Yes	Yes	Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix J: Heterogeneous results by campground popularity

This section explores heterogeneous welfare damages based on the popularity of a campground. We define popularity based on the average number of visitors per year for years in which it was open. For reference, Table A1 shows the most-visited campgrounds, many of which belong to high-profile national parks, such as Yosemite, Grand Canyon, and Rocky Mountain. The least popular tend to be small, local, or regional US Forest Service campgrounds. We rerun the main estimation but allow the smoke and travel cost coefficients to vary by the quartile of popularity. Given 999 campgrounds, each quartile contains approximately 250.

Figure J1 summarizes the point estimates for smoke responses, travel cost responses, and willingness to pay (WTP). Full results are displayed in Table J1. Across specifications, the magnitude for both the smoke and travel cost coefficients are lower at more popular campgrounds. These results suggest visitors are more willing to incur both higher travel costs and some environmental disamenity for highly desirable locations.

The translation of these responses to welfare impacts is less clear. WTP is estimated as the ratio of marginal disutility in smoke to that in expenditure (the smoke coefficient divided by the travel cost coefficient). Because WTP is a ratio, it could be either higher or lower given reductions in both the smoke parameter (the numerator) and the travel cost parameter (the denominator). Figure J1 shows that the reduction in the smoke parameter dominates, resulting in lower WTP at popular campgrounds. Table J1 confirms that WTP is lower at popular campgrounds across all specifications. In general, welfare damages tend to be largest for the middle two quartiles of popularity.

Figure J1: Point Estimates for Smoke Response, Travel Cost Response, and Willingness to Pay (WTP) by Quartile of Popularity Using Model (4) (visitors are less responsive to smoke and travel cost at more popular campgrounds)

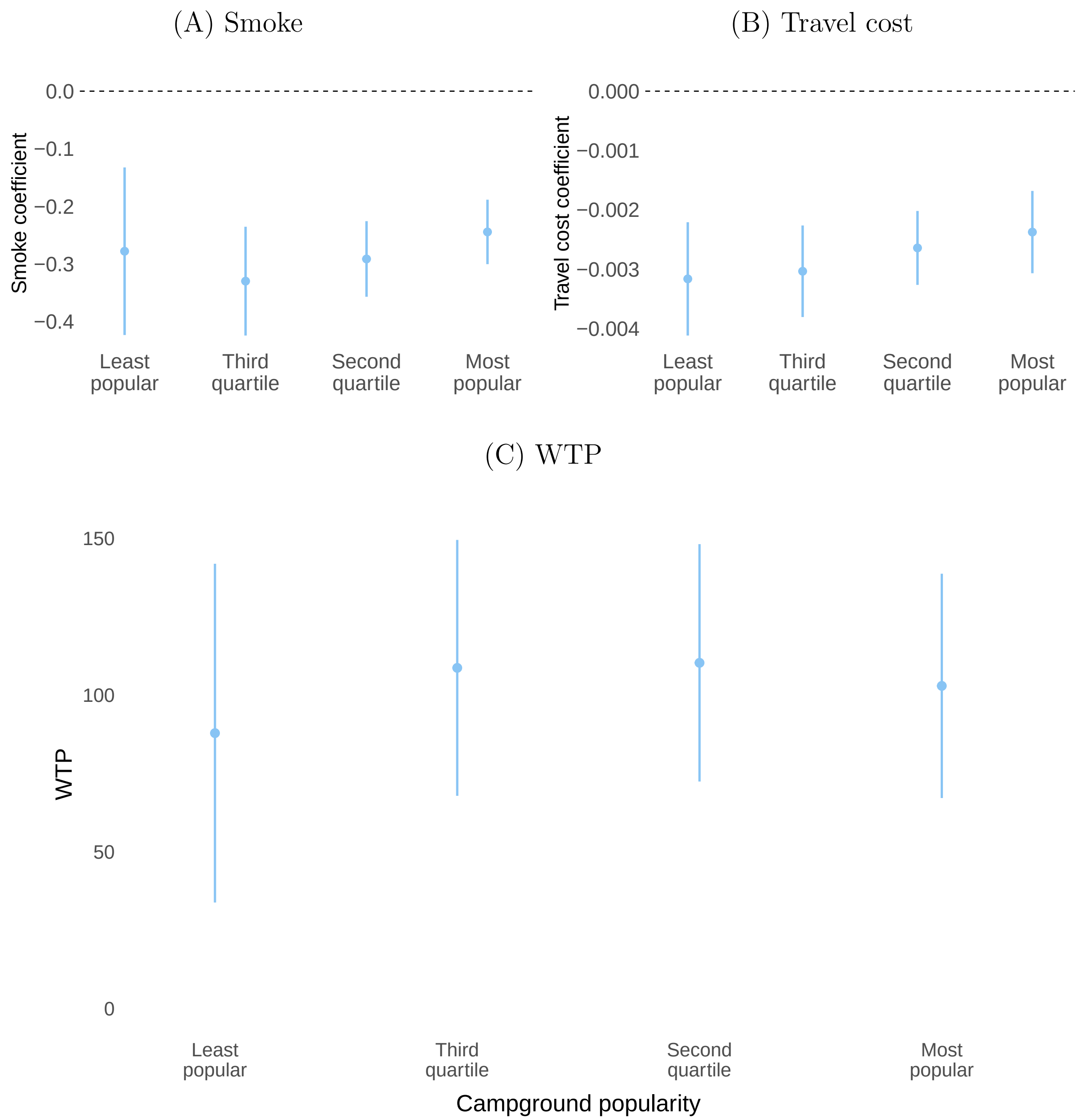


Table J1: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Heterogeneity by Campground Popularity

	(1)	(2)	(3)	(4)
Inv. distance to wildfire (km^{-1})	-11.0284**	-12.0844**	-11.9595**	-7.8196**
	(0.9225)	(2.4306)	(2.4448)	(0.8254)
High temp. (degrees C)	0.0187**	0.0289**	0.0293**	0.0307**
	(0.0044)	(0.0023)	(0.0023)	(0.0022)
Low temp. (degrees C)	-0.0012	-0.0204**	-0.0213**	-0.0252**
	(0.0055)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0046**	-0.0059**	-0.0061**	-0.0057**
	(0.0010)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0037	-0.0377**	-0.0387**	-0.0402**
	(0.0256)	(0.0120)	(0.0120)	(0.0122)
Smoke x first quartile (most popular)	-0.2208**	-0.2297**	-0.2035**	-0.2446**
	(0.0320)	(0.0345)	(0.0338)	(0.0286)
Smoke x second quartile	-0.2563**	-0.3296**	-0.3007**	-0.2915**
	(0.0425)	(0.0417)	(0.0407)	(0.0335)
Smoke x third quartile	-0.2364**	-0.3161**	-0.2889**	-0.3301**
	(0.0462)	(0.0482)	(0.0482)	(0.0482)
Smoke x fourth quartile (least popular)	-0.2488**	-0.3577**	-0.3457**	-0.2781**
	(0.0576)	(0.0673)	(0.0681)	(0.0743)
Travel cost x first quartile (most popular)	-0.0028**	-0.0023**	-0.0024**	-0.0024**
	(0.0004)	(0.0003)	(0.0003)	(0.0004)
Travel cost x second quartile	-0.0010*	-0.0027**	-0.0027**	-0.0026**
	(0.0005)	(0.0003)	(0.0003)	(0.0003)
Travel cost x third quartile	-0.0009	-0.0030**	-0.0030**	-0.0030**
	(0.0006)	(0.0004)	(0.0004)	(0.0004)
Travel cost x fourth quartile (least popular)	-0.0009	-0.0031**	-0.0031**	-0.0032**
	(0.0006)	(0.0005)	(0.0005)	(0.0005)
WTP: first quartile (most popular)	79.36**	98.63**	86.49**	102.87**
	(17.55)	(21.64)	(20.09)	(18.25)
WTP: second quartile	249.52	123.65**	112.36**	110.23**
	(136.76)	(22.99)	(21.6)	(19.31)
WTP: third quartile	253.55	106.16**	96.52**	108.62**
	(169.66)	(22.64)	(21.48)	(20.82)
WTP: fourth quartile (least popular)	272.18	116.18**	110.39**	87.8**
	(195.78)	(29.58)	(28.58)	(27.56)
N	2,723,034	2,691,655	2,691,655	2,688,739
Campground FE		Yes	Yes	
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix K: Total welfare estimate data construction

In Section 5, we report estimates for the total annual number of recreation visits affected by smoke in the west; we combine the Recreation.gov data with overall visitation data from various federal and state agencies. In particular, we use total visitation numbers from the National Park Service, US Forest Service, Bureau of Land Management, US Army Corps of Engineers, and National Association of State Park Directors. Each reports visitation at varying spatial and temporal levels. For example, the National Park Service reports visitation at a park by month level; the US Forest Service reports at a forest by year level; and the state parks report at a state by year level. For each data source, we aggregate the daily Recreation.gov data to the most relevant spatial and temporal scale to determine the proportion of visits affected by smoke. We then multiply this proportion by the total visitation data. In this section, we detail this process for each data source.

For the National Park Service, we use the agency's Annual Summary Reports.[33] This dataset reports total monthly visitation at all national parks, national monuments, national recreation areas, and other lands that the agency manages. In the western states, 27 national parks are included in the Recreation.gov dataset, and 82 are not. For those 27 parks, we determine each park's monthly proportion of campers that were smoke affected. We then multiply this proportion by each park's monthly visitation from the Annual Summary Reports to infer the total number of smoke-affected visits. For those 82 parks, we calculate a statewide proportion of smoke-affected campers in the data. We multiply these state by month proportions by each park's visitation levels in the Annual Summary Reports based on its location.

To estimate smoke-affected visits at national forests, we use the US Forest Service's National Visitor Use Monitoring (NVUM) Program.[34] These data report visitation at all National Forests at an annual level. In the west, 70 forests are included in the Recreation.gov dataset, and 8 are not. For those 70 forests, we calculate each forest's annual proportion of campers affected by smoke and multiply it by the corresponding annual visitation to-

[33] National Park Service. Annual Summary Report. `https://irma.nps.gov/STATS`.

[34] US Forest Service. National Visitor Use Monitoring Program. `https://www.fs.usda.gov/about-agency/nvum`.

tals in the NVUM data. For those eight forests, we use a statewide annual proportion of smoke-affected campers.

The Bureau of Land Management records visitation statistics as part of its Recreation Management Information System (RMIS).[35] We contacted the program administrator and received data on site by year visitation for all BLM sites.[36] Most visitation is not reservable, and a large portion is considered backcountry. Therefore, the Recreation.gov dataset contains very few BLM campgrounds. We thus combine annual state level proportions of smoke-affected campers from the Recreation.gov data with annual site visitation from the RMIS.

For sites managed by the US Army Corps of Engineers, we use data from its Value to the Nation (VTN) reports.[37] For the study period of 2008 to 2017, the agency only has one year of recreation data (2016). We treat this year as representative of typical annual visitation over the study period. For each site, we multiply the total number of visitors by the state level average of smoke-affected campers from the Recreation.gov data over all years.

Last, we estimate smoke impacts at state parks. We use visitation data from the National Association of State Park Directors, which was compiled by Smith et al. (2019). For these data, the unit of observation is a state by year. We again use annual state level proportions of smoke-affected campers from the Recreation.gov data multiplied by the NASPD data.

Having approximated total visitation, we multiply each agency's annual smoke-affected visits by the empirical estimate of per trip losses due to wildfire smoke. We estimate that more than 21.5 million recreation visits per year are affected by smoke in the west, with annual losses of $2.3 billion. For further discussion, see Section 5.

[35] Bureau of Land Management. Public Land Statistics. `https://www.blm.gov/about/data/public-land-statistics`.

[36] Ridenhour, L. and Leitzinger, K. Bureau of Land Management. Personal correspondence.

[37] US Army Corps of Engineers. Value to the Nation. `https://www.iwr.usace.army.mil/Missions/Value-to-the-Nation`.